J. Neurol. Neurosurg. Psychiat., 1955, **18**, 287.

A PRELIMINARY REPORT ON A PERCEPTUAL MAZE TEST SENSITIVE TO BRAIN DAMAGE

BY

ALICK ELITHORN

From the Neurological Research Unit of the Medical Research Council, The National Hospital, Queen Square, London

During an investigation into the clinical effects of prefrontal leucotomy, an attempt was made to design a simple psychological test procedure sensitive to organic damage to the frontal lobes. The preliminary results obtained with the perceptual maze test reported below suggest that it is sensitive both to cerebral damage in the frontal regions and to damage to the anterior temporal lobe. Because of the theoretical interest of this test further development and standardization studies will be undertaken. These involve a change in the format of the test material, and it seems appropriate to make now a preliminary communication describing the principles of the test and the clinical results obtained.

Description of the Test

The test material (Fig. 1) consists of 30 patterns each of which has a lattice background of paths or tracks. At a variable number of intersections dots are superimposed. In the original pattern the paths were white with red circles at the intersections and the background was black (Fig. 1a). In the new version the paths are interrupted lines and the dots are solid black circles (Fig. 1b, c, and d). In either version the subject is required to find in each pattern a pathway from bottom to top—or from top to bottom—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., it 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. In order that the subject may know when he has found a correct pathway, he is told the maximum number of dots which can be obtained in each pattern.

In a version which is being standardized, the individual items are printed together in a form suitable for administration as either a group or individual pencil and paper test. All the preliminary work, the results of which are reported here, was however carried out with the original type of pattern (Fig. 1a). These patterns were drawn up individually on separate sheets of paper. After preliminary trials 30 items were selected to form a scale of steadily increasing difficulty. Patterns were chosen on the basis of their difficulty relative to the size of the background trellis.

Administration of the Test

Each subject was shown a sample pattern (Fig. 1a) and was given the following instructions: "Here is a pattern; you have to find the path from bottom to top or top to bottom which passes through the most red circles. You must stick to the white paths and at each fork you can go either left or right but you must not double back." These points were illustrated on the demonstration pattern and repeated if the subject did not clearly understand them. This wording was not rigidly adhered to and if the subject's understanding was thereby aided concrete analogies, e.g., grass and paths, were used. The subject was then told that he would be shown other similar patterns and that for each of these he had in the same way to find a path which passed through the greatest number of circles. It was added that he would be told for each pattern how many circles it was possible to obtain. He was also told that he would be given plenty of time—one minute—for each pattern but that he should attempt to find the correct path as quickly as possible. In order that the test material could be re-used the subject was not allowed to draw on the patterns but was asked to show the examiner the correct path with a wooden stylus. The test was then given in accordance with the above instructions and a record was kept of the time—to the nearest five seconds—taken to solve each pattern. When the subject failed to find the correct answer within a minute he was shown the solution by the examiner and was then given the next pattern. The test was continued until the subject failed to solve four consecutive patterns. Because early experience had shown that the extra time did not add proportionally to the subject's ability to solve individual patterns the test was stopped when the subject failed to solve four consecutive patterns each within 30 seconds. This materially reduced the amount of time wasted due to subjects scoring lucky hits after they had already reached their ceiling. Until the test was stopped a record was kept of the time taken on all solutions achieved within the one minute time limit. The test was scored on the basis of one point for each pattern correctly solved with an additional credit of a further point if it was solved within 30 seconds.

Subjects

In its original form the test was administered to the following groups of subjects: Group A, patients who were examined with the maze both before and after leucotomy or temporal lobectomy; Group B, patients who were attending the leucotomy follow-up clinic, most of whom had been examined in the psychological department before and after operation but who were examined with the maze only after operation; Group C, other patients from the neurological and psychiatric wards of the hospital.

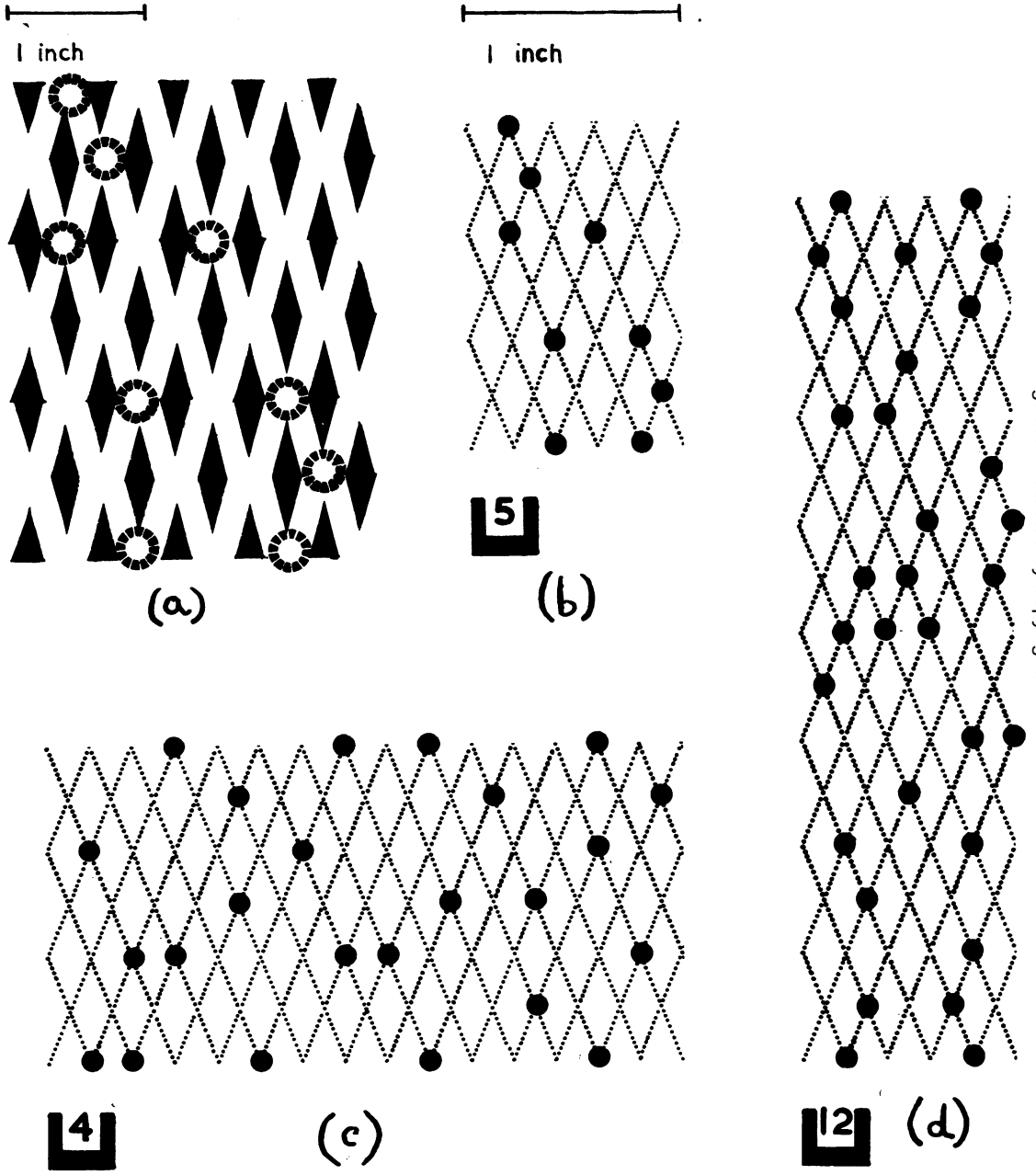


FIG. 1.—Reproductions of sample patterns: (a) Type of maze used for individual test reported in the present paper; in the original material the rings were red. (This particular pattern was used as a demonstration one.) (b) The same pattern in the modified format. (c) and (d) Sample patterns to illustrate that the complexity can be increased by enlargement in both the horizontal and vertical dimensions. The photographic reduction for the two types of pattern is different, the original sizes being as indicated.

Results

Group A : Temporal Lobectomy and Prefrontal Leucotomy.—Because a modified operation is now the treatment of choice in patients who come to leucotomy, 14 of the 20 subjects in this group had partial leucotomies. Only one patient who suffered from intolerable pain associated with an inoperable carcinoma had a standard leucotomy. The remaining five patients had an operation involving the partial removal of one temporal lobe. All the patients in this group were tested before operation and again not less than a fortnight post-operatively. Of the 14 patients subjected to a partial leucotomy 12 achieved a better score post-operatively. The mean gain for these 14 patients was $10{\cdot}1 \pm 2{\cdot}95$ points. This contrasts markedly with the results obtained with the five patients who underwent an operation on the temporal lobe. Of these, four performed less well post-operatively, the mean loss for all five subjects being $8{\cdot}6 \pm 6{\cdot}28$ points. The groups are small but the difference between them is statistically significant ($t = 3{\cdot}05$, $P < 0{\cdot}01$). Although four of the temporal lobe cases had a partial or complete quadrantic field defect their failure on the test did not appear to be due to this disability. Thus the one patient who had no field defect was also reported as having no visuospatial disability following the operation ; nevertheless his score was worse post-operatively. Conversely, the one temporal lobe patient whose performance was better post-operatively did have a left upper quadrantic field defect. This observation—that a small amount of damage to one temporal lobe disturbs performance on this test more than does a modified bilateral leucotomy—needs confirmation. It is, however, in accord with Milner's contention that damage to a temporal lobe causes impairment of high level visual skills (Milner, 1954).

Group B : Post-Leucotomy.—There were 36 patients in this group, all of whom were examined at least three months after leucotomy. Twenty-two of these patients had had a standard operation of the type described by Freeman and Watts (1942) ; 13 patients had had some form of partial leucotomy, and, of these, seven had had a modified operation by a lateral approach, five a rostral leucotomy, and one a cingulectomy ; one patient, who had had a standard leucotomy elsewhere and who was subsequently submitted to a partial operation here because it was considered that the first operation had failed, has been omitted. The mean score for the 22 patients who had had a full leucotomy was $18{\cdot}5 \pm 2{\cdot}26$ while that for the 13 patients who had a partial operation was $23{\cdot}8 \pm 3{\cdot}28$. This difference does not reach a satisfactory level of statistical confidence ($t = 1{\cdot}10$, $P > 0{\cdot}1$). The results are, however, suggestive, and examination of the case records of the patients showed that two of the patients in the partially leucotomized group were anomalous. One had proved to be suffering from a presenile dementia and the second had had a post-operative collapse attributed to a complicating cerebral thrombosis. If these two patients are excluded, then the remaining 11, who had only a partial operation, performed significantly better (mean score $26{\cdot}9 \pm 2{\cdot}76$) than those who had had a full leucotomy (mean score $18{\cdot}5 \pm 2{\cdot}26$, $t = 2{\cdot}31$, $P < 0{\cdot}05$). That this difference is not attributable to age differences between the groups has been shown by an analysis of covariance, and that it is not due to original differences in intellectual level has been similarly demonstrated on the 29 patients for whom data as to the pre-operative intellectual level were available. This observation that a standard leucotomy impairs the maze performance more than does a modified leucotomy does not conflict with the previous observation that patients retested after a modified leucotomy showed improved scores. This improvement would be partly due to practice effects and partly due to the reduction of emotional disturbance which followed the operation.

While the maze test was being administered to these patients attending the leucotomy follow-up clinic it was noticed that there was an apparent correlation between success on this test and successful post-operative adjustment. In order to test Porteus' claim (Porteus, 1952) that maze type tests may be sensitive to " social intelligence ", Mr. E. G. Glithero, Psychiatric Social Worker to the Department of Psychiatry, who was completely unaware of the results obtained with the test, was asked to assess the success these patients were making of their post-operative social rehabilitation. He very kindly undertook this rather difficult task and succeeded in producing for the group for whom psychological test results were available, a rank order of social adjustment. There is a significant relationship between this rank order of social adjustment and that of success on the maze. In assessing the strength of this relationship Kendall's tau has been used : ($tau = 0{\cdot}308$, $P < 0{\cdot}02$, $n = 31$). The value obtained for the relationship between pre-operative intelligence, as measured by the vocabulary score, and post-operative social adjustment is much smaller and not significant ($tau = 0{\cdot}156$, $P > 0{\cdot}05$).

Group C : Other Patients.—Unfortunately it did not prove possible to retest after the lapse of an appropriate interval the majority of patients in this group. Most of them were, however, also tested with the vocabulary and Koh's blocks sub-tests of

the Wechsler Bellevue scale (Wechsler, 1944), and the results obtained with these subjects have therefore been used together with the pre-operative findings obtained with the subjects of Group A to derive product moment correlations between the maze scores obtained and the two Wechsler sub-tests. The correlation between the maze score and the vocabulary score is $+0{\cdot}462$, $P < 0{\cdot}01$ ($n = 43$). That between the maze and Koh's blocks is $+0{\cdot}738$, $P > 0{\cdot}001$ ($n = 43$).

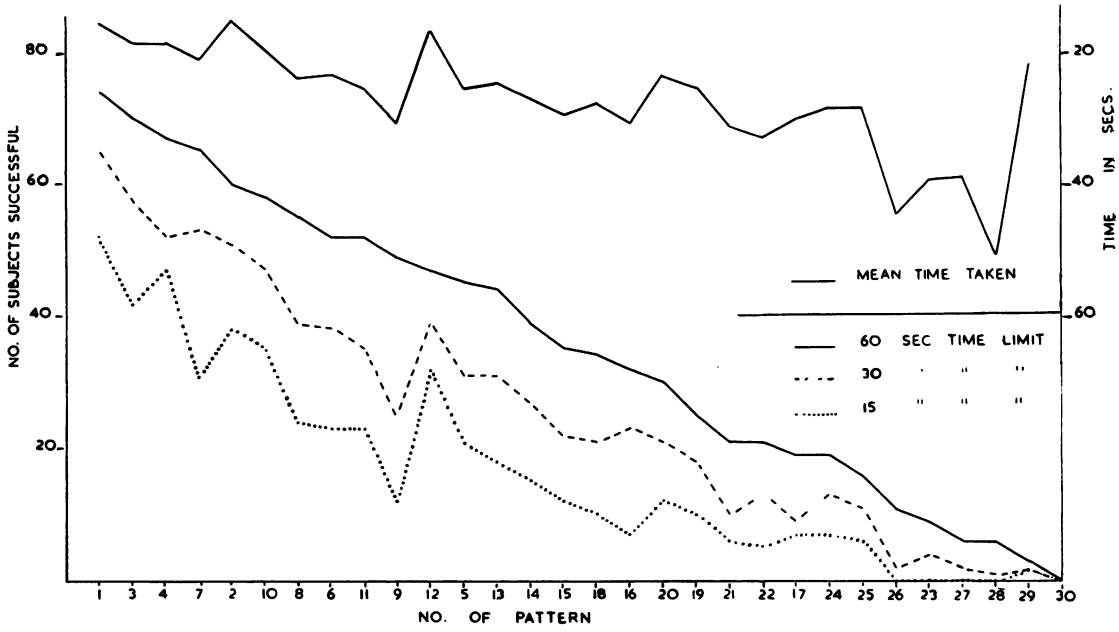
Performance on psychological tests sensitive to organic cerebral damage tends to deteriorate with increasing age and it seemed worth while to study the relationship of this variable to the maze scores. For this group of subjects the product moment correlation with age for the vocabulary scores was $-0{\cdot}040$ ($P > 0{\cdot}1$), for the Koh's blocks scores it was $-0{\cdot}575$ ($P < 0{\cdot}001$), and for the maze it was $-0{\cdot}500$ ($P < 0{\cdot}001$). The finding that the vocabulary score is virtually independent of age and that the Koh's block test is not is a well attested one (Jackson, 1955). The observation that the maze test performance with this group deteriorates with increasing age is in agreement with and supports the finding that this test is sensitive to organic deteriorative processes. It is not on the present material an independent observation: the group is not a normal sample and the frequency of organically determined psychological illnesses is greater amongst the more elderly patients. Similar considerations may have tended to produce a rather high correlation between the maze test scores and the Koh's block scores but it is worth observing that the partial correlation between these tests when the vocabulary score is held constant is still high ($r_{mk.v.} = +0{\cdot}589$). This suggests that the maze test may, like Koh's test, be a good test of spatial abilities.

The results obtained with Group C subjects together with those obtained from the patients in Group B and those obtained on first testing the patients in Group A have been used further in studying other characteristics of the maze patterns. The results for all the 87 subjects are presented graphically in Fig. 2. In this figure each co-ordinate refers to an individual pattern. These are arranged along the abscissa in order of increasing difficulty according to the number of successful solutions achieved during the 60 seconds allowed for each pattern. The diagram also shows the number of patterns which were achieved within the time limits of 30 seconds and 15 seconds respectively. It will be apparent that an increase in the amount of time allowed for each puzzle does not result in a proportionate increase in the number of subjects achieving a solution. For the test as a whole the following results were obtained: when 15 seconds was the time allowed, 497 items were solved; in 30 seconds, 762 items were solved; and when 60 seconds were allowed, 1,064 items were solved. It must be remembered that not all the subjects attempted the more difficult patterns as the test was discontinued after four consecutive failures. However, it was a commonly reported subjective experience that the correct pathway was seen either fairly rapidly or only with considerable difficulty. This all-or-none effect does seem to be characteristic of the test and may be related to another frequently reported tendency, namely, that subjects experienced difficulty in reconsidering paths that had already been rejected. In order to examine this point further the mean time for the successful solution of each item has been calculated. Again these figures have a limited value as there was an upper time limit of 60 seconds. The results have been plotted graphically (upper curve in Fig. 2) and it can be seen that the length of time taken to solve individual items increases very slowly when it is compared with increase in difficulty expressed as the number of subjects successfully solving the item within the time limit. This is particularly obvious over the median half of the test where the observations are likely to be the most reliable. Thus pattern No. 8 (seventh in order of increasing difficulty) was passed by 55 of the 87 subjects in a mean time of 24·2 seconds while pattern No. 25 (twenty-fourth in order of increasing difficulty) was passed by 16 subjects but in a mean time of 27·4 seconds. These 16 more able subjects solved pattern No. 8 in a mean time of 18·1 seconds.

Reliability of the Test

Although it has not proved possible to estimate the reliability of the test on a control group, evidence that this is high can be adduced from two independent observations. First the patients in Group A were re-examined on a second occasion after their operation. The interposition of an additional variable—type and extent of operation—will only dilute the reliability and it is likely that under more favourable circumstances the test-retest correlation would be higher than that observed ($r = 0{\cdot}813$, $P < 0{\cdot}001$, $n = 18$). Secondly the order in which the patterns were presented to the present group was determined by a preliminary pilot study with subjects the majority of whom were normal volunteers. If the correlation between the rank order of the items as determined in the pilot study and that observed in the present investigation is high it will indicate that the patterns are consistent in their

FIG. 2.—Relative difficulty of the different patterns. The patterns are arranged along the abscissa in an order of difficulty derived from the present investigation. The numbers of the patterns represent the order in which they were presented and this is the same as the order of difficulty determined in a pilot study. The ordinates for the three lower curves represent the number of successes in 15, 30, and 60 sec. respectively. The upper inverted curve shows the mean time in seconds taken to solve each test.



relative difficulty even when tested against samples from different populations. In Fig. 2 the patterns are arranged along the abscissa in an order of difficulty which is derived from the number of times each pattern was passed within the 60-second time limit during the present study. The numbers also represent the order of difficulty as determined in the pilot study. The relationship is clearly a close one (tau = 0·87).

Discussion

While the results so far obtained with this test are limited in range they indicate that the material is internally consistent and capable of giving repeatable results. That it may be sensitive to organic frontal lobe damage is indicated by the fact that it distinguishes between two groups of subjects who have been subjected to leucotomy operations of different extent. It appears to be unexpectedly sensitive to small lesions in the temporal region. It has also been shown that although the task is essentially a visuospatial scanning one, in which the need for formal intellectual processes is apparently absent, it correlates quite highly both with verbal and performance intelligence tests and with an independent assessment of post-operative social adjustment. Finally, test performance correlates negatively with increasing age.

The theoretical interest of this test lies in the fact that there is a high correlation between intellectual ability and test achievement, although once the instructions have been mastered the performance is essentially a perceptual skill and does not involve any formal logical processes. Thus it has been pointed out that successful solution of the task involves the ability to reject from the perceptual field irrelevant areas of the pattern and that once the correct pathway has been discovered it seems subjectively to dominate the perceptual field (Allan, 1954). It is of course true that the discovery of the correct path can, like the task of finding the ball in the field, be achieved by a rigorously pursued sequential search. Nevertheless most subjects feel that a solution is achieved more rapidly by a perceptual scanning of the material. As the difficulty of the patterns increases so does the tendency for this type of scanning behaviour to give way to the more sequential type of trial and error searching. It is suggested that superior performance on this type of maze is correlated with the ability to maintain a perceptual set in the face of material of steadily increasing complexity. In other words, the individual's performance reflects the size of the perceptual unit with which he is able to deal.

In this connexion it is interesting to note that Hick (1952), applying information theory to per-

ceptual tasks, has suggested that the rate of gain of information which can be achieved in perceptual tasks is in order of five bits a second. Although the maze test consists of a complex of binary situations, i.e., the presence or absence of a dot at sites which are distributed in a regular manner over the perceptual field, it cannot easily in its present form be treated in terms of information theory. Nevertheless, preliminary inspection suggests that this type of display may enable information to be taken up at a considerably greater rate than that observed in the more restricted situations studied by Hick (1952); this suggestion, though tentative, would be in accord with the observations of Klemmer and Frick (1953) on the assimilation of information from dot and matrix patterns. A further observation of considerable theoretical interest is that performance on the multiple choice maze seems to be easily impaired by emotional disturbance. The improvement which occurred in patients who responded satisfactorily to a modified leucotomy is almost certainly not entirely due to practice effects but partly to the reduction in emotional disorganization which generally accompanies a successful operation. It is in fact a common observation that emotional distress does interfere, often seriously, with the exhibition of intelligent activity in ordinary social situations. Further, Raven has pointed out (personal communication) that some reports that the reliability of his matrices is not always exceptionally high may only be a reflection of the fact that this test is sensitive to day-to-day fluctuations in intellectual efficiency.

That many valid and useful performance tests of intellectual ability measure the range of acquired intellectual acts and the speed with which these sets can be applied rather than the available integrative ability is a point not often emphasized. Cattell in distinguishing between these two types of intellectual performance called the first "crystallized" and the second "fluid" (Cattell, 1943). This author points out that it is fluid intellectual ability which declines most rapidly with age and which is most susceptible to brain injury. Similar considerations apply to the impairment of test performance by emotional disturbance. Halstead (1947), using physiological discriminative tests with a high "G" loading, found that his "impairment factor" was present almost as markedly before leucotomy as after. It may be as much this sensitivity to emotion as a practice effect which causes scores on the present test to improve after modified leucotomy. In addition to this inherent susceptibility to emotion of tests requiring "fluid G" the multiple choice maze appears to be in itself a neuroticizing situation. Because of the arrangement of the material it is not possible, in the absence of gross disorientation, to make an incorrect answer and consider it to be correct: the puzzles appear deceptively easy and subjectively a feeling of frustration can appear and increase rapidly when the subject has searched thoroughly, as he thinks, a matrix, without finding the correct path.

Summary

A multiple choice maze suitable for use as a clinical psychological test is described.

The preliminary results reported suggest that the test is sensitive to cerebral damage localized to the temporal or frontal lobes.

The performance of subjects on this test correlates significantly with estimations of their intellectual abilities derived from their performance on the vocabulary and Koh's block sub-tests of the Wechsler intelligence scale and with an independent estimate of success in post-leucotomy social adjustment. It correlates negatively with increasing age.

It is suggested that the success achieved in dealing with the test material may reflect the subject's ability to maintain a perceptual set in the face of material of steadily increasing complexity.

My thanks are due to the physicians and surgeons of the National Hospital who kindly allowed me to examine patients under their care, and to Professor O. L. Zangwill and his colleagues for access to and permission to use the results of tests undertaken in the Department of Psychology. To Dr. E. A. Carmichael, Dr. E. T. O. Slater, and Professor O. L. Zangwill I am indebted for criticism, advice, and continual support. I am also particularly grateful for stimulating discussion and help given by Miss M. D. Allan and Mr. M. F. Piercy. Miss M. A. Crosskey and Mrs. C. Lawrence assisted with the art work and the administration of the tests and undertook the major part of the statistical calculations. To Dr. P. Armitage I am grateful for statistical advice.

The material for this test, the N.R.U. Multiple Choice Maze, is not, at the time of writing, available in a standardized form suitable for routine use. Should it be required for research work, both the individual version and a form suitable for group administration can be obtained through the Medical Research Council's Neurological Research Unit at the National Hospital, Queen Square, London, W.C.1.

References

Allan, M. D. (1954). Personal communication.
Cattell, R. B. (1943). *Psychol. Bull.*, **40**, 153.
Freeman, W., and Watts, J. W. (1942). *Psychosurgery*. Thomas, Springfield, Illinois.
Halstead, W. C. (1947). *Brain and Intelligence: A Quantitative Study of the Frontal Lobes*. Univ. Chicago Press, Chicago.
Hick, W. E. (1952). *Quart. J. exp. Psychol.*, **4**, 11.
Jackson, C. V. (1955). *J. clin. Psychol.*, **11**, 137.
Klemmer, E. T., and Frick, F. C. (1953). *J. exp. Psychol.*, **45**, 15.
Milner, B. (1954). *Psychol. Bull.*, **51**, 42.
Porteus, S. D. (1952). *The Porteus Maze Test Manual*. Harrap, London.
Wechsler, D. (1944). *The Measurement of Adult Intelligence*, 3rd ed. Williams and Wilkins, Baltimore.