MENTAL TESTS IN SENILE DEMENTIA.

By H. HALSTEAD, M.A.,

Mill Hill Emergency Hospital Psychological Department.

(Received March 8, 1944.)

IN a previous article (1) the writer gave the results of a scale of 25 mental tests on a group of 20 senile patients. When that investigation was completed the scale was administered to a second group of seniles (Group B), who were considered to be in a more advanced state of dementia than the first group.

The objects of the second investigation were:

1. To see which of the tests discriminated clearly between the two groups.

2. To find out what kind of results would accrue from a sample of patients who were grossly deteriorated and in the closing phase of life.

3. To collect information which might help to determine the presence or absence of organic impairment in younger people. In this connection the hypothesis was made that in senile dementia deterioration of brain tissue would be diffuse, and that results on most tests would be poor. At this level neurotic overlay might be expected to be comparatively slight, and an uncomplicated record of organic loss could be obtained.

The first part of the paper sets out quantitative findings. The second part consists of behavioural data with notes on technique and mental deterioration.

SUBJECTS.

These were 18 unselected senile dements, all uncertified. Their ages ranged from 70 to 83, the average being 76·1, almost the same as for Group A.

RESULTS.

A. *Tests.*—The scores of both groups on each of the 25 tests were set out side by side, and compared with each other on four different indices, viz. (1) the critical ratio of the score differences; (2) the percentage of loss suffered by Group B in relation to Group A; (3) the percentage of overlap in scores between the two groups; and (4) the percentage of subjects in each group who failed completely on any of the tests. Of these indices the most reliable is the percentage of loss in productivity shown by Group B, as the differences are obtained, not from average scores, but from the total weight of scores of both groups on each test. In order to arrive at the best assessment all four indices were taken into account, and Table I shows the discriminative value of the tests under the headings "Good," "Fair," and "Poor." With such small numbers (38 patients in all) score distributions are unreliable, but the tests starred gave a fairly normal distribution.

Both groups found little difficulty in Recognizing Photographs, Naming Months, and Packing Cubes in Box.

Knox Cubes, Porteous Mazes, Tapping, and Recall of Paragraph were difficult for all the subjects.

On the whole finer differences were obtained on the timed tests (scored in seconds), which, besides producing a wider dispersion of scores, measure efficiency (2).

When the tests were ranked in order of difficulty for both groups the two lists were different. After eliminating those tests with similar rankings the discrepancies showed that Group B experienced *relatively* more difficulty in Reversing Months, Repeating Digits Forward, Recalling a Paragraph, Naming Colours in Reverse, Arithmetic, and Porteous Mazes. In relation to Group A three tests were much higher in rank order, i.e. relatively easier; these were Naming Colours, Tapping, and Tracing.

TABLE I.—*Discriminative Value of the 25 Tests Between the Two Groups of Senile Patients.*

Good Discrimination.

* Picture Memory.
Reading Speed.
* Digits, Forward and Reversed.
* Colour Naming, Reversed.
Recall of Paragraph; immediate delayed.
Naming Months, Reversed.

Fair Discrimination.

Block Designs (Kent).
Vocabulary.
* Knox Cubes.
* Naming Birds.
* Arithmetic.
Porteous Mazes.
Recognizing Photographs.
* Sentence Memory.

Poor Discrimination.

Naming Months.
Naming Colours.
Tapping.
* Naming Animals.
Tracing Tests (2).
* Writing Tests (2).
Packing Cubes in Box.

N.B.—Tests marked * show a fairly normal distribution over the 38 subjects.

TABLE II.—*Approximate Mental Ages of Group* B *Patients on* 13 *Tests.*

Test.	Mental age.
Vocabulary	10·5
Naming Months	9·5
Arithmetic	8·0
Reading Words	7·0
Repeating Digits, Forward	7·0
,, ,, Backward	7·0
Sentence Memory, rote	7·0
Knox Cubes	6·5
Block Designs, easy series	6·0
Naming Months, Backward	5·0
Recall of Paragraph, immediate	5·0
,, ,, delayed	5·0
Porteous Mazes	5·0
Average mental age, 13 tests	6·8

Table II sets out the average mental ages achieved by Group B on 13 of the tests. The average age for these tests is 6·8 (median = 7 ± 1·5 years). The highest mental age, as one might expect, is on the Vocabulary Test—mental age 10·5. Between this and the worst tests there is a gap of 5·5 years. As far as test performance generally is concerned the average demented senile in this group was, therefore, little better than that of a child of seven.

B. *Subjects.*—Dealing now with patients instead of tests, the following differences appear between the two groups of seniles:

1. On the full scale (25 tests) there was a large gap between the median ranks of the two groups. The median of Group A was 13·77 ± 6·39, and that of Group B 23·58 ± 5·73, out of 38 ranks. The critical ratio of this difference is 3·72, which is decidedly significant. (P = 99 to 1.) The overlap in ranks between the groups was 20 per cent., i.e. eight cases were misplaced according to the test results.

2. The average loss of production from Group A to Group B on the full scale was 36 per cent., which is considerable in view of the advanced age of the patients in *both* groups.

3. The average percentage of failures on the tests was 22 per cent. in Group B, as against only 5 per cent. for Group A. The largest discrepancies in this respect were in Reversing Months, Recalling a Paragraph (both immediately and after a delay), Naming Colours in Reverse, and Arithmetic. There were no failures in either group in Repeating Digits Forward, Tapping, and Packing Cubes.

4. As far as variation of scores on the 25 tests is concerned, there was little to choose between the two groups, but there was a greater tendency in Group B for variability of performance to increase with age and with lowering of the general level of ability.

5. Group B showed a steeper drop in mean scores with increasing age ("r" = −·55 as against "r" = −·27 in Group A).

C. *Test-profiles.*—When the test-profiles of the 38 patients were examined it was found that no two were alike. This is to be expected, since, even at this stage of life, there is a dispersion of scores on each test, but especially on the timed tests.

In comparison with the adult male population generally the senile scores are low in every test, but individual differences are still present. It has been shown that "intelligence" ceases to develop earlier in people of low than of high intelligence, and the present study indicates that, in general, the decline of abilities is more rapid in the former, if the vocabulary score is used as an index of former level. But deterioration of brain tissue, though it may be diffuse in seniles, gives rise to different behaviour patterns, and it was not easy to predict how any patient would perform on a given test, unless he were so deteriorated that he could not comprehend the situation. The pattern and tempo of decay may be determined at conception, and modified by the life experience of the individual. In younger people, where variation of performance is greater, prediction is more difficult, and certainly less exact. Further, the concern of the psychologist is usually with intellectual impairment, which is only a part of the total picture. In any senile patient one could not deduce with confidence from the test records the state of emotional deterioration or lapse in personal habits.

Even in cases of known cerebral arteriosclerosis there are still considerable variations in score profiles, so that, although they may be grouped diagnostically, they cannot be so grouped in test performance. In Group B there were six cases, and their profiles differ widely from each other. Their rankings on the scale were 19, 21, 23, 26·5, 37 and 38, i.e. all in the lower 50 per cent. The average of these is not significantly different from the average of Group B, and their average age differs only by six months from that of the same group. The two poorest cases out of the 38 were both arteriosclerotic and both aged 80.

Observations on Tests.

It is an accepted fact that in senile decay memory is almost always affected, especially memory for recent events. For this reason the test scale was heavily weighted with memory tests both of rote and substance memory, in the visual and auditory spheres.

Visual Memory.

Experimentally it has been shown that recognition is one of the easiest tasks of memory, and might, therefore, be expected to show the slowest decline. In the test of Recognizing Photographs, however, a test which, in the form used, was found to be too easy for most adults to whom it was applied, the demented patients in Group B were poor, even in comparison with Group A. Only one patient obtained a full score, which is awarded for recognizing four pictures singly and then all four in order (8 points). The mean score was 6 points ± 1·5. In this group there was a good deal of indecision, confusion, and slowness, and in some cases the patient became haphazard or immobilized. In addition to the poor memory, retroactive inhibition and perseveration both played a part in confusing the patients. On the Picture Memory test, a card of ten familiar objects exposed for 30 seconds, Group B patients were markedly inferior. Three failed completely. The median score

was only 4 ± 2 objects recalled. The highest score was 7. Few seemed, on being questioned, to have been able to conjure up an image of the card once it was removed. In the worst cases the time of exposure was much too short for the patients to attend to more than a few items, and results showed that these were mostly on the top line. (There were two lines of five.) Retention was so fleeting that even these were lost almost as soon as the card was being taken from sight. Increasing the time made little difference to the results. The poor score reflects the slow comprehension, brief attention, and short retentive powers of the demented seniles.

Another visual test, Knox Cubes, in which the subject has to tap, from memory, a given number sequence, immediately after its demonstration by the examiner, showed up the narrow and fleeting visual span. The median score of 4 ± 1·5 indicated an ability to reproduce only a simple sequence of four taps. The impression faded so quickly that unless the patient made his attempt at once he had to guess. As the series increased in length or complexity memory of the early part faded during the demonstration, and led to confusion. This test, unless done by a "counting" method, which allows mental rehearsal, involves a kinaesthetic element as well as a visual. On inquiry it appeared that most of the patients used the visual kinaesthetic method.

Auditory Memory.

Rote or verbatim memory is usually easier than substance memory, of meaningful material. It measures attention and retentive powers, but not the active digestion involved in substance memory. Two rote tests were included in the scale, viz. Digits and Sentence Memory. Wechsler says (3): "Rote memory, more than any other capacity, seems to be one of those abilities of which a certain absolute minimum is required, but excess of which seemingly contributes relatively little to the capacities of the individual as a whole. The memory span for digits has the great merit of quickly indicating whether an individual has that relative minimum." The median scores for Group B patients were 5 ± 0·5 forward, and 3 ± 1 in reverse, a performance which, in younger people, would lead one to suspect feeble-mindedness, according to Wechsler.

In the test of Sentence Memory (rote) the percentage of loss shown by Group B patients was rather less than in digit span. Only half of the Babcock series was used, viz. ten sentences graded in difficulty. In this short series there was too sudden a jump from an average of 10 to an average of 20 words. In the shorter sentences Group B averaged 87 per cent. correct, i.e. only about 8 words verbatim (Group A, 96 per cent.). Doubling the quantity reduced this to 57 per cent., or an average of 10 words (Group A, 71 per cent., or 14 words). As the sentences became longer omissions and inventions increased, and in Group B jargon appeared. (In Group A it appeared only once, in the worst case.) Once the limit of verbatim memory has been reached the subject has to rely on substance memory and guessing. Apparently the saturation point had not been reached in Group A, as there was a tendency to get a higher percentage of words correct as the sentences increased in length. In Group B the limit had about been reached ("r" = + 0·2). If the verbatim memory for speech is so short "spatially" and temporally, it follows that the demented senile will have difficulty in following a conversation and will miss many points altogether. This is one reason why he makes so many inconsequential remarks. All test instructions had to be given slowly, and in many cases repeated.

Substance Memory was measured by getting the subjects to reproduce a 46-word paragraph, containing 20 "ideas," from memory. Few people can reproduce this amount of material verbatim, and, in practice, there is usually a combination of verbatim and interpretive recall. Most seniles fare badly on this test, but in Group B the median number of ideas reproduced was *nil*, and 11 patients failed completely. The best effort produced only 6 ideas out of 20. The total amount of material recalled immediately by all the patients in Group B was only 6·7 per cent. of the possible (Group A, 24·5 per cent.). After a delay of a few minutes the number of total failures in Group B rose to 15. In general, as far as the demented senile is concerned, spoken words seem "to go in one ear and out of the other." The total production of all Group B patients in the delayed recall test was a mere 4·2 per cent. of the possible (Group A, 20·3 per cent.).

Fluency.

It was known that senile dements were slow in their reactions, especially in thinking. Even old associations are affected, though less than recent ones. Four short " fluency " tests were given, viz. Naming Animals and Birds, Reading simple Monosyllabic Words, and Naming Months. The first three have a time limit of one minute ; the fourth is timed, but the limit depends upon the subject.

In Naming Animals and Birds Group B showed a loss of 33 per cent. relative to Group A. The median scores were only $7 \pm 5{\cdot}5$ and $6 \pm 2{\cdot}5$ in the minute respectively. The former group showed considerable blocking (pauses) both in length and frequency, and many repetitions, only a few of which were recognized as such by the patients. Perseveration from one test to the other occurred only in Group B, and then only in a few cases.

In Reading Words and Naming Months the demented seniles were very slow, and there was often difficulty in shifting from one word to the next, with some slurring and repetitions. A few of them tried to read the words as if they were prose.

The Vocabulary Test also deals with old associations, but this has been given a section to itself.

Psychomotor Tests.

The tests in this category comprise Block Designs, Tracing (drawing lines through dotted patterns), Writing, and Packing Cubes. All these tests were within the capacity of most of the Group B patients. The only failures recorded were on the Block Designs as the patterns became more difficult. The series used here was that designed by G. H. Kent (4) for younger children ; there are only four designs and no dual colour blocks are used. The chief features observed in this group of tests were slowness and tremors.

Vocabulary.

The Bellevue Vocabulary Test was used in this investigation because of the graded scoring and the good distribution of scores found in adults. It was preferred to a selective type of test because it evokes material which can be assessed both quantitatively and qualitatively. The Shipley Vocabulary Test was tried at first, but the form of the test was irksome to the patients ; they tended to confuse one line with another.

In the Inventive Test the subject is required to show, not only his acquaintance with a word, but also how it is used and what it means. There are many grades of definition, and a skilled examiner can tell in a few minutes whether he is dealing with a person of high or of low verbal intelligence.

Radosaka-Strzemeka (5) has given the genetic steps in the ability to formulate definitions, and adults become more or less fixated at one of these levels. In both senile groups all levels were manifest.

The person of good previous intelligence had not lost the ability to define *per genus* and *differentiae*, a process involving classification and abstraction. One senile dement said, not quite accurately, that a donkey was " An animal of the quagga species." Even at this level there are finer grades according to the nearness or remoteness of the class concept or the adequacy of the words used in defining.

Definition by synonym is usually indicative of good level and economy of words, e.g. " To bet " for Gamble.

A median level response is definition by use or properties, e.g. " You wear it," " You ride on it," etc. Here, again, there are qualitative differences, e.g. for bacon many patients said " You eat it." One patient preferred " You have it for breakfast "—a rather more concrete definition.

Definition by description of qualities seems, in adults, to be at a lower level, though Terman (6) places it higher than the last named, genetically, e.g. " It has four legs," " It is round and red," etc.

Definition by giving an instance is at a lower level still. One patient said for " Nuisance," " Well I am a nuisance." (Why ?) " Because I get in the way."

One of the more primitive forms is the ostensive definition or pointing out an instance, e.g. " That is a nail in the wall," or " This is an apple." If the patient has no example available he may be nonplussed.

Sometimes the subject has a vague idea of what a word means, but is unable to frame a definition, often saying that it is on the tip of his tongue. This happens fairly often with uncommon words, though it is not unusual to find a subject in difficulty in trying to define a common word such as cushion or nail. Inability to define words which are "known" is often indicative of seriously impaired function, since the breakdown involves old and familiar associations. One deteriorated patient, in trying to do the Vocabulary Test, either gave up the attempt to define a word or merely repeated the word to be defined, e.g. "A donkey, well it's a donkey, isn't it?"

If a person is unable to make any score on a Vocabulary Test he is in a poor mental condition, i.e. if he is not suffering from a specific defect such as aphasia. The lowest ranking patient of both senile groups scored only one mark on the Vocabulary Test despite variations of technique. His sole definition was "Bacon—Oh, they don't eat it do they?"

For ordinary clinical use the graded scoring of the Vocabulary Test is quite adequate, but for more careful assessments it is a good plan to appraise the definitions on a qualitative basis.

Testing Technique.

Taken individually, each of the tests in the scale is quite short, and the apparatus simple. Most of it consists of material typed on cards, one test to a card. These are easily carried, and form a useful adjunct to a clinical interview.

The test session started with interesting visual tests, such as Recognizing Photographs, and Picture Recall. These usually break the ice fairly well. After this the tests were varied, i.e. visual and auditory material alternating, with psychomotor tests interspersed. Fairly easy tests completed the session. Most of the tests are timed, and a stop-watch is a necessity if one is to compare any performance with the norms. The procedure throughout was standardized without being rigid. In addition to the test scores careful notes were made about each patient's behaviour, including remarks and conversation, as well as any oddities of performance. This is the clinical method of testing which teaches the examiner a good deal about a patient.

Mental Deterioration.

The investigation of senile patients brought out many points worth following up in other cases of organic impairment.

In the first place it showed the rapid decline of recent memory span generally and of substance memory particularly in old people. Every test in the scale demonstrated impairment, but it was possible to rank the tests in order of difficulty to bring out the different rates of decline (1). The clinical tester should have a series of age curves covering the greater part of life on a variety of tests. These would show graphically the mean scores and standard deviation (or similar measures) at each age. This would not be an easy task, but it would provide a quantitative criterion for clinical use.

Secondly, it showed the impossibility of assessing deterioration by short-cut methods. A variety of tests must be used to find the patient's strong and weak points. But this gives rise to difficulty as, despite the known differential decline rates, individual test profiles differ widely, even in old age. This raises the question whether score patterns can be related to clinical syndromes. Wechsler (3) has attempted this task, but differential "diagnoses" based on his score profiles are not sufficiently clear; there is overlapping between different groups. It would seem impossible to obtain profiles which are mutually exclusive, i.e. which apply to one and only one type of case. One remedy seems to be the accumulation of "signs" as in a clinical diagnosis, or, say, in the Rorschach test.

Thirdly, it indicated the advisability of paying special attention to certain tests which are found to be awkward for most organic cases, extracting from them all possible information. Many researchers have used this method, chiefly in the study of categorical thinking applied to different kinds of material. Some of the memory tests in the present scale will repay more intensive investigation.

Finally, at the risk of being trite, one might mention again the difficulty experienced by so many organic cases in shifting attention from one activity to another.

This difficulty affects some activities more than others, but in each, the patient, having assumed a mental set, seems tied to it (" Stimulus-bond "), and the succeeding activities suffer in different degrees from an occasional lapse into some preceding test, to a complete inability to start a new one. This was frequently noted in the seniles, especially in Group B, both in and between tests.

In trying to determine the presence and extent of organic impairment the investigator may need to try several methods, and will benefit by trying them on different pathological types. He will also need behaviour norms for different ages. As Gesell (8) says, *à propos* of children, "A developmental diagnosis is essentially an appraisal of the maturity of the nervous system with the aid of behaviour norms."

Summary.

1. A scale of 25 mental tests was given to a group of 18 demented senile patients (Group B) aged 70 to 83. The results were compared with those of a previous group (Group A) who were in a less advanced stage of dementia. All the tests differentiated between the two groups. Table I shows the various discriminative values of the tests.

2. Mental ages for Group B on 13 of the tests are shown in Table II. The median of these is $7 \pm 1{\cdot}5$ years.

3. When the average of all the test ranks of Group B was compared with that of Group A the difference was decidedly significant and there was only a 20 per cent. overlap. Other criteria, such as total production and percentage of failures on tests, also showed significant differences. There is little difference in " scatter " between the two groups of patients, but the more demented group show a steeper drop in scores as their ages increase.

4. There is little in common among the score profiles of all the 38 patients, nor among those of six cases diagnosed as cerebral arteriosclerotics.

5. Observations on test performances appear under the headings of Visual Memory, Auditory Memory, Fluency, Psychomotor Tests, and Vocabulary.

6. The senile investigation as a whole brought out a number of points relevant to the study of organic impairment; these are discussed.

I wish to thank the Ella Sachs Plotz Foundation, and Dr. V. Korenchevsky, of Oxford, for their kind assistance; Dr. Roberts, of St. Francis and Dulwich Hospitals, and Dr. Turnbull, of Tooting Bec Hospital, for allowing me to see cases under their care; Dr. Aubrey Lewis for his encouragement; also the Hon. W. S. Maclay and Dr. Louis Minski for permission to undertake the investigation.

References.

(1) Halstead, H. (1943), *J. Ment. Sci.*, **89**, 376–377.
(2) Babcock, H. (1941), *Time and the Mind.* Sci-Art Publishers, Cambridge, Mass.
(3) Wechsler, D. (1942), *Measurement of Adult Intelligence.* Williams, Wilkins Co., N.Y.
(4) Kent, G. H. (1934), *J. App. Psychol.*, **18**, 578.
(5) Radosaka-Strzemeka (1930), *Bull. international de la société scientifique de pédagogie.* Warsaw.
(6) Terman, L. M., and Merrill, M. A. (1916), *The Measurement of Intelligence.* Harrap.
(7) Klopfer, B., and Kelly, D. McG. (1942), *The Rorschach Technique.* World Book Co., N.Y.
(8) Gesell, A., and Amatruda, C. S. (1941), *Developmental Diagnosis.* Harper Bros.