THE MEMORY FUNCTION

I. A Factorial Study of Fifteen Clinical Tests

H. J. EYSENCK, Ph. D.,[1] and H. HALSTEAD, B. A.
Psychological Department, Mill Hill Emergency Hospital
Mill Hill, England

1. Introduction

Tests of the memory function are of obvious importance to psychiatry in view of the fact that memory defects are considered symptomatic of various syndromes, such as senile dementia, the Korsakoff syndrome and brain injuries. A variety of so-called memory tests have therefore been developed for use by the clinician, which include such tasks as repeating a story, repeating digits or letters, describing pictures from memory, and many others of a similar kind.

Memory tests are also of importance to systematic psychology, where work on learning, retention, recall and recognition has made great strides since the early days of Ebbinghaus. In particular, interest has centered around the question of the exact status of "memory" within the total field of cognitive mental ability(33). The method most frequently used to attack this problem has of necessity been that of factor analysis.

There is considerable agreement among workers in this field that "memory" represents a group-factor within the wider field of cognitive ability; Anastasi(2, 3), Brener(6), Bryan(7), Carlson(9), Garrett(14), Holzinger *et al.*(17, 18, 19), Kelley(23), Spearman(34), Thurstone(35, 36), and Woodrow (40) have reported studies to this effect.

While such general agreement is welcome in a science which often shows extreme divergence of findings, closer scrutiny of the actual tests used and the methods of analysis employed, must make us chary of accepting the conclusions as established fact. For instance, Carlson(9) finds that "a general factor, common to all the tests in [his] battery, is by far the most important single factor involved in the memory of words." But his interpretation of this factor as one of "memory ability," while certainly inviting, does not by any means follow directly from his data; quite possibly we are dealing here simply with general mental ability (Spearman's "g" factor), expressed through the medium of memory tests.

A similar criticism is applicable to Brener's study of memory span(6). He finds high correlations between 13 different tests, administered under conditions of visual or auditory presentation. He accordingly concludes that there is "ability common to all of the tests. . . . . The general factor is tentatively described as the memory span factor." Again, the possibility remains that we are merely dealing with general intelligence, not with a special memory factor.

When it is further realized that of the tests used in all the studies mentioned, a large number were group tests, quite different in kind from the usual clinical tests, and that their interpretation as "memory" tests is often open to doubt, we must come to the conclusion that the existence of a group factor of "memory," linking together all the tests involving learning, retention and recall or recognition, is rather more questionable than appeared at first.

In the more clinical field of memory assessment, also, it appears at first sight that there is much agreement with regard to the existence of the "memory ability" discussed by the factor analysis. Wells and Martin(39), Simmins(32), and Landis and Rechetnick(24) found evidence of memory defects in senile dements, G.P.I. patients and epileptics, while finding little impairment in the "functional" groups tested. Others, such as Rawlings(29), Foster(13), Babcock(5), and Shakow, Doecart and Goldman(31) have found impairment in functional disorders also. Achilles(1), Liljencrants(25), Moore (27), and Wylie(41) have attempted to show differences between the psychoses by comparing test results obtained by different

[1] With the support of the Rockefeller Foundation.

procedures (method of recall *vs.* method of recognition).

While these various studies are of great interest, they fail to prove the existence of a separate "memory" ability for two reasons. In the first place, general intelligence has not always been partialled out, so that the relative contributions of intelligence and the hypothetical "memory" factor cannot be determined. In the second place, Hull(21, 22) and Wechsler(38) have shown reason to believe that the impairment found in various clinical groups is due to failure at the learning stage rather than to failure in retention; when material was learned to a certain standard by controls and patients of various clinical groups (paretics, schizophrenics, etc.), Hull actually found the patients superior in ability to reproduce the learned material after one week(22).

The real difficulty in accepting such evidence as has been presented, either factorially or clinically, in favour of the existence of a separate factor of "memory ability" lies in the impossibility of accepting the usual type of memory test as anything but a test of learning ability. Memory involves the three stages of learning, retention and recall or recognition; the usual tests of memory hardly ever proceed beyond the first stage, penalizing the tests so heavily for imperfect learning that possible differences in retention or recall are obliterated. Only when learning is eliminated as a factor, as in the Hull experiment, can retention be tested; but such elimination is hardly ever attempted in the usual type of memory experiment.

In the present study an attempt was made to investigate the factors active in 15 memory tests of the usual type, and to determine to what extent success in these tests was determined by intelligence. We hope to follow this up by an experiment in which a similar determination of factorial patterning is carried out for what we would regard as more genuine tests of "memory," *viz.* tests of retention in which the learning aspect has been equalized.

2. The Experiment

The subjects used in this investigation were all male neurotic army patients under treatment at the Mill Hill and Sutton Emergency Hospitals. They were selected on the basis of their scores on the progressive matrices test of intelligence(28), which is given routinely to all incoming patients. Previous work had shown that neurotics do not differ to any considerable extent from normal subjects with respect to intelligence, or to the degree of reliability of their test scores (11, 12, 15). Twelve patients were taken from each of the five selection grades as defined by the test, making 60 patients in all.[2]

In addition to the intelligence test, which was given as a group test, 15 memory tests were given to all the patients. These memory tests were given as individual tests, and are described below in some detail. As it seemed important to avoid fatigue in the patients, the memory tests were given in two groups, on separate days. On each occasion, the patient was tested for approximately 35 minutes; wide individual differences in testing time were encountered, however.

The series of memory tests comprised 10 auditory and 5 visual tests. They were given in the order as described below;

1. *Digits forward.*
2. *Digits backward.*

   These two tests were taken from the Bellevue scale(37), the procedure used there was taken over without modification.
3. *Cowboy story.*

   This story was taken from Henderson and Gillespie's textbook(16). It was read aloud by the experimenter to the patient; immediately afterwards the patient was given a "quiz" consisting of 16 questions relating to the story; one mark was allowed for each question answered correctly.
4. *Photograph recognition.*

   This test consisted of 12 black and white head-and-shoulder photographs of American men appearing in an issue of *Time*. They were uniform in size and all were full-face. Three of the photo-

[2] It is customary to divide subjects into five grades on the progressive matrices test: grade I, including the highest 10%; grade II, the next 20%; grade III, the middle 40%; grade IV, the next 20%; and grade V, the lowest 10%.

graphs were placed in front of the patient for 30 seconds; they were then returned to the pack, which was shuffled. The patient was asked to pick out the three photographs he had just seen, and place them face upwards in the same order on the table. This procedure was repeated with 4, 5, 6, 7 and 8 photographs. The cards were numbered on the back, and each series presented followed the same number-sequence as was used in test 1. A score of 2 was given for each card correctly placed in order. If a card was correctly chosen but placed out of order in the sequence, a score of 1 was given.

5. *Paired associates.*

A selection of 10 word-pairs was made from Babcock's scale(4). Each pair of words was read aloud with an interval of 2 seconds between pairs. After a short pause the experimenter read the first word of each pair, and the subject had to say the associate word. Each correct response scored 1 point.

6. *Cube imitation (Knox).*

Four cubes are placed in front of the patient. With a fifth cube the experimenter taps the four cubes according to a number sequence, and the patient has to do the same immediately afterwards. The number-series increase in difficulty. One point is awarded for each correct sequence. The procedure was taken without modification from Cattell(10).

7. *Word recognition.*

Twenty words were selected from the Babcock test(4); 10 of these were read out at the rate of one per second. Then the complete list was read, and the patient had to indicate in response to each word whether or not it was one of the original ten. Each correct reply scored 1 point.

8. *Donkey story.*

This story was taken from Henderson and Gillespie's text-book(16). After the story was read aloud by the experimenter, the patient was asked to reproduce as much of it as possible in his own words. The story was divided into 41 "ideas," each scoring 1 point.

9. *Designs recognition.*

This test was similar to test 4 in all respects, except that designs, taken from Hornung's book(20), took the place of photographs.

10. *Letters forward.*

11. *Letters backward.*

These two tests were similar to tests 1 and 2 in all respects, except that letters took the place of digits.

12. *Word Memory.*

Six series of monosyllabic words, consisting of 3, 4, 5, 6, 7 and 8 words respectively, were read aloud by the experimenter, and the patient was asked to repeat them. Each word reproduced in correct order was scored 2 points; a score of 1 was given to a word correctly reproduced but not in order. This test was taken from(8).

13. *Sentence memory.*

This test was taken from Babcock(4). Ten sentences of increasing length were read aloud by the experimenter, and the patient was asked to repeat each one immediately. The scoring was different from Babcock's, each word correctly reproduced being allotted 1 point.

14. *Design reproduction.*

This test, taken from Rey(30), falls into two parts. In the first, the patient copies a complex design from a card. Test 15 is then interposed. Finally the patient is required to reproduce the design from memory. Our scoring procedure followed that of Rey. This is the only test of incidental memory in our battery.

15. *Picture memory.*

A card with 16 objects drawn in ink was put in front of the patient for 30 seconds. It was then removed, and the patient had to enumerate as many of these objects as possible. Each object correctly remembered scores 1 point.

3. Results

Tetrachoric correlations were run between the 15 memory tests used, and are reported in Table 1, together with Spearman rank-correlations (corrected) between scores on these 15 tests and scores on the intelligence test. (The distribution of scores on the memory tests was sufficiently normal to permit use of the tetrachoric co-efficient. The distribution of scores on the intelligence test being rectangular rather than normal because of the selection of subjects, we thought it more appropriate to use Spearman's rho, rather than the tetrachoric coefficient which presupposes a normal distribution of scores.)

TABLE 1

Test No.:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1.	..	.74	.58	.78	.46	.70	.44	.87	.75	.82	.70	.67	.78	.71	.41	.71
2.		..	.67	.71	.51	.81	.35	.78	.68	.70	.90	.71	.82	.82	.41	.89
3.			..	.72	.71	.63	.54	.84	.82	.94	.70	.60	.67	.69	.42	.78
4.				..	.67	.81	.82	.95	.98	.82	.82	.84	.72	.92	.71	.87
5.					..	.67	.60	.81	.71	.49	.67	.59	.51	.67	.55	.67
6.						..	.59	.86	.84	.64	.81	.78	.77	.87	.71	.81
7.							..	.74	.82	.78	.67	.56	.43	.69	.74	.67
8.								..	.94	.86	.84	.80	.87	.89	.70	.81
9.									..	.80	.78	.89	.76	.87	.74	.84
10.										..	.95	.88	.57	.88	.42	.88
11.											..	.92	.78	.93	.63	.96
12.												..	.69	.89	.59	.84
13.													..	.80	.59	.72
14.														..	.63	.92
15.															..	.63
16.																..

It will be seen that all the correlations are positive, and that most of them are comparatively high. The correlations between the various memory tests and the intelligence test are seen to range from +0.63 to +0.96.

A factor analysis was carried out on the correlations reported in Table 1, the results of which are given in Table 2. The method of analysis used was Burt's summation method with successive approximations. One general factor, accounting for 74% of the variance, was found sufficient to account for all the correlations within the limits of the probable error; none of the residuals was statistically significant. Also given in Table 2 are the means, standard deviations, and coefficients of variation of the various tests.

TABLE 2

	Factor-Saturation	Mean Score	S.D.	Coefficient of variation
1. Digits forward	.78	6.9	1.3	19
2. Digits backward	.81	4.9	1.5	31
3. Cowboy story	.80	10.0	3.8	38
4. Photo recognition	.95	51.0	15.9	31
5. Paired associates	.72	4.7	2.7	57
6. Cube imitation	.88	7.2	2.4	33
7. Word recognition	.73	16.7	3.3	20
8. Donkey story	.98	18.9	7.7	41
9. Design recognition	.96	48.2	15.2	32
10. Letters forward	.89	6.2	1.4	23
11. Letters backward	.94	4.5	1.1	24
12. Word memory	.88	44.0	10.4	24
13. Picture memory	.81	10.6	3.2	30
14. Sentence memory	.95	17.3	4.4	25
15. Design reproduction	.68	27.5	10.9	40
16. Intelligence	.94	...	...	..
Variance:	.74	Average:		31

The nature of the general factor found cannot be decided by mere inspection. Seeing that 15 out of 16 tests used were "memory" tests of the usual type, it would be tempting to identify the resulting factor as a "memory" factor. A statistical examination shows that such a procedure would be definitely fallacious.

Seeing that all the memory tests correlate with the intelligence test, the question that has to be answered is: With the influence of intelligence excluded, does anything remain which is common to all or nearly all of the memory tests? Two methods of answering that question suggest themselves.

If the factor which is being measured by the intelligence test were identical with the factor which is being measured by the 15 memory tests, then we should be able to treat the correlations of each of these memory tests with the intelligence test, corrected for attenuation,[3] as factor saturations; it would follow that the product matrix calculated by multiplying these correlations should be approximately identical with the actually observed correlations between the 15 memory tests. This is indeed the case, as can be shown by subtracting the product matrix thus derived from the original matrix of intercorrelations: the residual matrix contains only one correlation more than twice the size of the S.E. of a zero correlation, and in a table of this size, we would expect at least one correlation of this size by chance. The average of the residual correlations, calculated regardless of sign, is only +0.09. Thus we can account for all the observed correlations without the necessity of postulating a memory group factor in addition to a general factor of intelligence.

A second method of proof also follows from the general theory of factor analysis (*cf.* Burt, C., "Factors of the Mind," or Thurstone, L. L. "Vectors of Mind.") The matrix test is known to be an almost pure test of "g" or general intelligence; in other words, its validity as a test of "g" is a function of its reliability, being equal to the square root of its reliability. Now if the battery of memory tests measured "g" purely, and without any admixture of a memory group factor, then this total battery of 15 tests should be a good approximation to a pure "g" estimate. That being so, the saturation of the matrix test for the general factor found in the analysis of the total battery of 16 tests should be equal to the square root of its reliability. This is indeed the case, and we may argue back from this finding that our assumption regarding the nature of the battery of 15 tests was justified, *viz.* that these tests did not contain any factor over and above "g."

Our two methods of analysis, then, agree in showing no evidence of a group factor of memory. Because of the small number of subjects used, and because many other tests of memory apart from those used could be devised and might give different results, we do not claim to have disproved the possible existence of such a factor. We do claim, however, that if such a factor exists its contribution to the total score must be small as compared with the contribution of general intelligence, and that no assessment of memory ability is possible by the use of the tests examined.

4. Discussion

There are various points in connection with this research which require discussion. The first point relates to the selection of patients. In many ways it would have been preferable to have dealt with patients distributed normally with regard to intelligence, rather than with a sample giving a rectilinear distribution, such as ours. The main reason for choosing our sample in such a way that the five selection grades were represented equally was simply that we wanted to get tentative norms for the various tests used, for each of the selection grades separately; had we selected a normally distributed sample there would have been too few patients in grades I and V to give us norms of any value whatever.

One effect of this choice of subjects is of course that all the correlations between the various tests are rather higher than they would otherwise have been; this is an inevitable consequence of the greater heterogeneity of the sample. But there is no reason to suppose that this greater heterogeneity would tend to mask a factor of "memory ability," had such a factor been present; thus our main conclusion remains unaffected by the abnormal distribution of our subjects with regard to intelligence.

A second point relates to the question of

[3] That is to say, attenuation arising through the lack of complete reliability in the intelligence test. The unreliability of the memory tests affects equally their correlations with each other and with intelligence. The reliability of the intelligence test was determined by retesting a sample of 60 patients whose scores on the matrix test were distributed similarly to those of our experimental population, and was found to be +0.89 ± .03.

whether results derived from testing neurotics can have any general validity. There seems little reason for assuming that cognitive organization is essentially different for neurotics, as compared with the normal population; certainly our experience with other cognitive tests does not support such a view (11, 12, 15). While we believe that a repetition of the experiment with larger samples of both tests and subjects, normal and abnormal, would throw further light on this vexed question of "memory," we also believe that our results are sufficiently definite to allow us to say that the practice of testing memory by the usual clinical tests is unsound, and that these tests measure intelligence rather than memory. While it is possible that a slight memory factor may have escaped through the none-too-fine meshes of our probable errors, its contribution to the total score on any of the memory tests examined would not be large enough in any individual case to outweigh the much more considerable contribution of the general cognitive factor.

5. Summary and Conclusion

Fifteen individual memory tests and one intelligence test were given to sixty male neurotic army patients at Mill Hill and Sutton Emergency Hospitals. The scores on these 16 tests were intercorrelated, and the resulting table factor-analyzed. The following results were found:

1. All the memory tests correlated positively with the intelligence test, the correlations varying from +0.63 to +0.96.

2. One general factor accounted for all the correlations in the table within the limits of the probable error; this general factor accounted for 74% of the variance.

3. It was shown by two separate methods that the general factor found could be equated with general intelligence, or "g," and that all the correlations in the table could be accounted for without the necessity of postulating a "memory" factor.

4. The conclusion was drawn that the ability involved in the clinical tests of "memory" studied in this research was identical with that involved in the intelligence test used, and that therefore it was misleading to accept scores on these various tests as estimates of a person's "memory" ability.

We are indebted to the superintendents of Mill Hill and Sutton Emergency Hospitals for permission to use their clinical material.

BIBLIOGRAPHY

1. Achilles, E. M. Experimental studies in recall and recognition. Arch. Psychol., **6**, No. 44, 1920.

2. Anastasi, A. A group factor in immediate memory. Arch. Psychol., **18**, No. 120, 1930.

3. Anastasi, A. Further studies on the memory factor. Arch. Psychol., **22**, No. 142, 1932.

4. Babcock, H. An experiment in the measurement of mental deterioration. Arch. Psychol., **18**, No. 117, 1930.

5. Babcock, H. Dementia præcox, a psychological study. Lancaster, Pa.: Science Press, 1933.

6. Brener, R. An experimental investigation of memory span. J. Exp. Psychol., **26**: 467-482, 1940.

7. Bryan, A. I. Organization of memory in young children. Arch. Psychol., **24**, No. 162, 1934.

8. Bronner, A., Healy, W., *et al.* A manual of Individual Mental Tests and Testing. 1927. Little, Brown & Co., Boston.

9. Carlson, H. B. Factor analysis of memory ability. J. Exp. Psychol., **21**: 477-492, 1937.

10. Cattell, R. B. Guide to Mental Testing. London: University of London Press, 1936.

11. Eysenck, H. J. The effect of incentives on neurotics, and the variability of neurotics as compared with normals. Brit. J. Med. Psychol., **20**: 100-103, 1944.

12. Eysenck, H. J. Neurosis and intelligence. Lancet, Sep. 18, 1943, 362.

13. Foster, J. C. Significant responses in certain memory tests. J. Appl. Psychol., **4**: 142-154, 1920.

14. Garrett, H. E. The relation of tests of memory and learning to each other and to general intelligence in a highly selected adult group. J. Educ. Psychol., **19**: 601-613, 1928.

15. Halstead, H. An analysis of the matrix (progressive matrices test results on 700 neurotics (military subjects, and a comparison with the Shipley Vocabulary test. J. Ment. Sci., **89**: 202-215, 1943.

16. Henderson, D. K., and Gillespie, R. D. Textbook of psychiatry.

17. Holzinger, K. J. Relationships between three multiple orthogonal factors and four bifactors. J. Educ. Psychol., **29**: 513-519, 1938.

18. Holzinger, K. J., and Harman H. H. Comparison of two factorial analyses. Psychometrika, **3**: 45-60, 1938.

19. Holzinger, K. J., and Swineford, F. A study in factor analysis: the stability of a bi-factor solution. Suppl. Educ. Monogr., No. 48, Univ. of Chicago, 1939.

20. Hornung, C. P. Handbook of Designs and Devices. 1932. New York: Harper & Bros.

21. Hull, C. L. The formation and retention of associations among the insane. Am. J. Psychol., **28**: 419-435, 1917.

22. Hull, C. L. Quantitative aspects of the evolution of concepts. An experimental study. Psychol. Mon., **28**: 123, 1920.

23. Kelley, T. L. Crossroads in the mind of man. Stanford Univ. California, 1928.

24. Landis, C., and Rechetnick, J. Changes in psychological function in paresis: I. Psychiat. Quart., **8**: 693-698, 1934.

25. Liljencrants, J. Memory defects in the organic psychoses. Psychol. Mon., **32**, No. 143, 1922.

26. McNemar, Q. The Revision of the Stanford-Binet Scale; an Analysis of the Standardization Data. Boston: Hougton-Mifflin, 1942.

27. Moore, T. V. The correlation between memory and perception in the presence of diffuse cortical degeneration. Psychol. Mon., **27**, No. 120, 1919.

28. Raven, J. C. Standardization of Progressive Matrices, 1938. Brit. J. Med. Psychol., **19**: 137-150, 1941.

29. Rawlings, E. The intellectual status of patients with paranoid dementia præcox; its relation to the organic brain changes. Arch. Neur. Psychiat., **5**: 283-295, 1921.

30. Rey, A. L'examen psychologique. Archives de Psychol., **28**: 112-164, 1941.

31. Shakow, D., Doecart, M. B., and Goldman, R. The memory function in psychotics. Dis. Nerv. Syst., **2**, 43-48, 1941.

32. Simmins, C. Studies in experimental psychiatry: IV. Deterioration of "g" in psychotic patients. J. Ment. Sci., **79**: 704-734, 1933.

33. Spearman, C. The Abilities of Man. New York: Macmillan, 1927.

34. Spearman, C. Thurstone's work re-worked. J. Educ. Psychol., **30**: 1-16, 1939.

35. Thurstone, L. L. The perceptual factor. Psychometrika, **3**: 1-17, 1938.

36. Thurstone, L. L. Primary mental abilities. Psychometr, Monogr. No. 1, 1938.

37. Wechsler D. The Meausurement of Adult Intelligence. Baltimore: Williams & Williams Co., 1939.

38. Wechsler, D. A study of retention in Korsakoff's psychosis New York: Dissertation, Columbia University. 1917. Quoted in Hunt, McV., (Ed.), Personality and the Behavior Disorders. 1944. New York: Ronald Press.

39. Wells, F. L., and Martin, H. A. A method of memory examination suitable for psychotic cases. Am. J. Psychiat., **8**: 243-258, 1923.

40. Woodrow, H. The common factors in fifty-two mental tests. Psychometrika, **4**: 99-108, 1939.

41. Wylie, M. An experimental study of recognition and recall in abnormal mental cases. Psychol. Mon., **39**, No. 180, 1930.