MEMORY SPAN TESTS

By H. J. Humpstone, Ph.D.

University of Pennsylvania.

The memory span test consists of a presentation of series of stimuli with the immediate reproduction of the series by the subject. The stimuli may be digits, syllables, words, sentences, paragraphs, diagrams, objects; the presentation may be made auditorily, visually, articulatory, graphically or in combinations of two or more of these methods. Usually the stimuli are presented either auditorily or visually, but the subject may and in most cases does re-enforce the stimulus by repeating the elements or naming the objects, so that it is next to impossible to get a purely auditory or visual stimulation. Nor is this very desirable, because in every day life from childhood we habitually use some such methods of combining impressions, and the results obtained from these tests are more significant if the methods of apprehension and reproduction are like those ordinarily used. Memory span test was first proposed and used by Jacobs, an English philosopher, and the results were reported by him in 1887. He used digits and syllables for stimuli, but discarded the latter in favor of the digits. In this country Smedley first used the test on a considerable scale in a study of the school children in Chicago to test methods of teaching spelling. He reported his results in 1900. Since that time the test in some form or other has been used by many workers, and when Whipple published his Manual of Mental and Physical Tests in 1910 he compiled a bibliography of fifty-two monographs and articles dealing with what he calls tests of rote memory, all of which are memory span tests in some form or other.

The assumption that this test is a test of memory has marred most of the experiments, and especially the interpretations which have been put upon results obtained from them. Jacobs in proposing the test called his study, "Experiments in Prehension." We think that a definite congenital ability exists, which is measured in the performance of this test. This is the ability to distribute the attention over a series of discrete perceptions, and is shown by their

immediate reproduction. In other words, it is the ability to grasp a number of different elements in one operation of attention. Perhaps "prehension" is a good word for this ability, but like so many other terms is used in other senses. We are using the term "Associability" and subsuming it under the general heading of imagination. The subject must have enough imageability to get perceptions of the stimuli, but the test does not measure imageability either as to types employed, or the extent to which they are involved in the performance. Memory is not tested. The images must be retained long enough but this process is only a sort of after image, and the period of its duration is so brief that the results do not furnish any criterion by which to judge of retentiveness. In the memory span test proper there is only one presentation of the stimulus, and this single presentation cannot be considered as teaching or training. Therefore, the only two aspects of that particular form of imagination called memory, namely, trainability and retentiveness, cannot be measured nor even estimated from the result of this test. These statements will be shown to be true from experimental data later on.

For several years we have been giving in connection with our first course in psychology, memory span tests, employing as stimuli, digits, letters, nonsense syllables, short words, sentences, paragraphs. These are presented auditorily with graphic reproduction. The results obtained in this way using digits as stimuli are presented herewith. In giving this test, series of digits are read at the rate of one per second without rhythm or grouping. The experiment is started with series of four digits, and the series are increased in length until series of twelve are reached. Two series of each length are given and the class is always informed of the different number of digits in the series about to be given. In preparing these series great care must be taken to have the digits really discrete stimuli. Reuther in "Psychologische Studien" in 1906 formulated rules for the construction of test series of digits. The most important are: "Do not place any two digits in their natural order, nor in their order reversed. Do not use sequences that suggest historical dates." These rules are violated in the Terman revision of the Binet test, and in the recently published series in the magazines, making such series worthless for testing purposes in comparison with the results of carefully prepared experiments. The students hold their pencils in their hands in a vertical position in order to avoid the tendency to make the figures in the air at least. As soon as each series is dictated the students write it as nearly as they can. In scoring the papers the number of digits in the longest series reproduced exactly as read is taken as the score. Using this method of procedure with several

succeeding classes we have the following table, including results from 476 men and women.

Digits	Number of students	Per cent of whole number
5	13	5.9
6	40	8.0
7	87	13.5
8	106	22.3
9	99	20.8
10	64	18.3
11	38	8.4
12	28	2.7

Mode=8. Median=8. Mean=8.5.
70.5% score above 7 digits.

Using the same method of procedure we employ as stimuli common words of three letters arranged in series of three or more words, two series of each length. Care is taken in arranging these series so as to avoid as far as possible familiar combinations and associations. We obtain this next table including results from 360 men and women.

Words	Number of students	Per cent of whole number
3	2	0.5
4	17	4.7
5	142	39.5
6	107	29.7
7	75	20.8
8	17	4.8

Mode=5. Median=6. Mean=5.8.
55.3% score above 5 words.

We have tried different sets of sentences none of which are entirely satisfactory. Those of the Terman scale are too easy and too absurd for our students. This year we have tried a new series ranging in length from ten to fifty syllables. There are two sentences of each length, one intended to introduce a concrete visual image, the other to avoid visual imagery so far as possible. Fifty syllables are too many to get exactly. The highest number anyone reproduced correctly is forty. The sentences which avoid concrete visual

imagery are reproduced correctly oftener than the others. The table includes results from ninety-one men and women.

Syllables	Number of students	Per cent of whole number
20	9	9.9
25	10	11.0
30	60	65.9
35	11	12.0
40	1	1.1

Mode=30. Median=30. Mean=29.2.
79% score more than 25.

For ideas we have also tried different material. Among others we used the well-known paragraph of the Binet series beginning "Many opinions have been given on the value of life." This is easily scored on the basis of seven fundamental ideas, and gives the following results from 135 men and women:

Ideas	Number of students	Per cent of whole number
1	1	0.7
2	3	2.2
3	13	9.7
4	31	22.9
5	30	22.2
6	31	22.9
7	26	19.2

No mode. Median=5. Mean=5.9.
87.2% score more than 3.

This shows that the test is too easy for our students, although it is included in the superior adult tests of the Terman Revision.

To complete this series of tests a real memory experiment is given consisting of a series of twenty sentences, each made up of a fact, a name, and a date. These are taught to the class by reading them over three times. After a period of forty-five minutes which is filled by a discussion of the former experiments and their relation to memory, the experimenter reads the facts, and the class writes the name and date. This experiment permits of several variations, and reproduction may be repeated at intervals to test retentiveness. In a class of eighty-nine men and women the scores range from six to

thirty-six. There is no distinct mode. The median is 20, the mean is 18.6. The names are remembered better than the dates. Probably the correlation would be higher with more verbal material used instead of the dates. Comparatively little has been done with this particular investigation, but it promises very interesting results.

In studying the results of these tests to gain light on their meaning we have found the correlation between the results with the different groups and with different tests. Using Pearson's formula the following are some of the results obtained. Any value of "r" less than .30 is held to show that the same factors are not operative or in other words that different abilities are involved in the tests. Any values over .30 show some tendency or ability operative in both tests, while values over .50 are considered to be very significant.

Digits and words, r = +0.52.

Digits and sentences, r = +0.37.

Sentences and ideas, r = +0.37.

Sentences and memory, r = +0.12.

Combined scores of digits, words, sentences and ideas with the memory experiment, r = +0.09.

These correlations are a strong indication that the discrete stimuli test some different ability from any of the other kinds of material; that this ability is not memory; that a person may be deficient in one ability and have enough of some other to make some compensation. We have found in our work with children that deficient associability usually is accompanied by difficulty in learning. Children who are advanced in school (under age for grade) have longer memory spans than those at age for grade, and they again longer than those over age for grade. This test taken in connection with a performance test is likely to be the most significant of all the tests we employ for diagnosis.