1957

A Study of Memory Span with the Bernreuter Memory Test

Joseph Michael Angileri
Loyola University Chicago

A STUDY OF MEMORY SPAN WITH THE BERNREUTER MEMORY TEST

by

Joseph Michael Angileri

A Thesis Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Master of Arts

June

1957

LIFE

Joseph Michael Angileri was born in Detroit, Michigan, August 6, 1932.

He was graduated from University of Detroit High School, Detroit, Michigan, June, 1950, and from the University of Detroit, June, 1954, with degree of Bachelor of Philosophy.

He began his graduate studies at Loyola University in September, 1954. He was employed by the Loyola Center for Guidance and Psychological Service from February, 1955 to August, 1956, in the Bureau of Child Study of the Chicago Board of Education from September, 1956 to April, 1957. He is presently employed by the St. Vincent and Sarah Fisher Home for Children in Farmington, Michigan.

ACKNOWLEDGEMENTS

Very special thanks are due to Robert G. Bernreuter, Frank J. Kobler, my advisor, and to the thesis readers; without their kindness and assistance this project could not have been completed.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
	Statement of the Problem	
II.	REVIEW OF THE LITERATURE	3
III.	THE PROCEDURE.	19
IV.	ANALYSIS OF RESULTS AND DISCUSSION	24
	Responses, corrections needed in the spans, and reasons for corrections	
V.	SUMMARY AND CONCLUSION	31
BIBLIOGRAPHY		35
APPENDIX		37

LIST OF TABLES

TABLE		Page
I.	EBBINGHAUS' FINDINGS ON HOW THE TIME FOR LEARNING SYLLABLES INCREASES WITH EACH TRIAL	4
II.	OBERLY'S STUDY OF PROGRESS OF MEMORY SPAN RESULTS	7
III.	TEST-RETEST DIFFERENCES OF THE FOUR SUB-GROUPS STUDIED BY MALDOWSKY	14
IV.	THE MEAN TEST-RETEST DIFFERENCES IN THE MALDOWSKY STUDY	14
V.	STANDARD DEVIATIONS FOR NUMBERS, LETTERS, AND NONSENSE SYLLABLES	27
VI.	THE RELIABILITY COEFFICIENTS USING THE SPEARMAN-BROWN FORMULA FOR TEST LENGTHENING	28

CHAPTER I

INTRODUCTION: STATEMENT OF THE PROBLEM

Clinical practice offers certain anomalies in the testing of auditory rote memory. Subjects are often inconsistent in levels of rote memory for digit spans, words and sentences; yet many clinicians in practice evaluate rote memory in terms of digit span alone.

The memory span for digits forward test has a reliability only in the sixties. This means that a digit forward span test does not measure memory span with a degree of reliability. This type of test also suffers from the fact that it does not lend itself to refinements of measurement. In the Terman-Merrill standardization, for example, an increase in one digit (from 4 to 5) is equivalent to two and one-half years of development, and an additional increase in one digit (from 5 to 6) is equivalent to three years of development. Thus, if a child with a "true" memory span of five years were, on one occasion, to repeat one fewer than this, and on another occasion to repeat one more than this, the clinician would be forced to report an increase in memory span from four and one-half to ten years. Such changes are to be expected frequently so long as the test has a reliability in the sixties.

Bernreuter is now undertaking the development of memory span testing materials, and is studying the reliability of new testing materials. He concluded that a memory span test can be

constructed using two forms of a number test and two forms of a letter test, which is reliable and sufficiently discriminatory to be used in individual diagnosis.

The hypothesis is that there is no significant difference between the performance of sixth grade students and Bernreuter's fourth grade students. The sixth grade students will perform better on number tests and letter tests than they will on syllable tests. Also, the combining of number tests and letter tests will further increase the reliability of a memory span test. Syllables will have little or no effect on the reliability of a memory span test.

CHAPTER II

REVIEW OF THE LITERATURE

Before discussing the results of this investigation, one must first look to others who were interested in memory span.

Men such as David Hartley and Thomas Brown were interested in memory and contributed to the existing theories concerning memory. Both were concerned with memory as an associative process, but did little to attempt to measure this process.

James Mills, in his Analysis of the Phenomena of the Human Mind, published in 1829, reduced all the laws of association to contiguity in experience. He stated that memory in the form of recognition is an idea or image of an object plus awareness of earlier experiences.

It was not until the late eighteen hundreds that new materials and procedures for measuring memory span were devised. In the monograph, Memory, Ebbinghaus brought to light the results of his five years of self-study. Ebbinghaus invented and used nonsense syllables in a study of his own memory processes. He used nonsense syllables - combinations of three letters, such as rof, guk, bap, rik, tid. Among the many areas he investigated were those of meaning and rhythm. He found that meaningless material was nine times as difficult to learn as meaningful. In the area of rhythm, he found that no one type of rhythm is necessarily best for all individuals, because individual differences are great.

He also found that if a memory task is made longer its learning time increases. Ebbinghaus discovered that he could recall seven or eight nonsense syllables by reading them once. To learn ten syllables required thirteen readings. The required time for each new item and for the whole list increased, but not proportionately. His findings appear in the following table.

TABLE I[1]

Ebbinghaus' Findings on how the Time For Learning Syllables Increases With each Trial

Length of List	Number of Readings	Time for Lists in Seconds	Average Time per Syllable in Seconds
7	1	3	.4
10	13	52	5.2
12	17	82	6.8
16	30	196	12.0
24	44	422	17.6
36	55	792	22.0

These are just a few of the problems investigated by Ebbinghaus in the field of memory. The other aspects of memory investigated by him are overlearning, spaced and unspaced learning, whole versus part methods, and the curve of forgetting.

---

[1]Herman Ebbinghaus, Memory, Trans. by Henry Ruzer and Clara E. Bussenus, (New York, 1913).

In 1871, T. L. Bolton conducted a research project to measure the memory span for digits. Fifteen hundred subjects, most of whom ranged in age from nine years to fifteen years, cooperated in the study. A few of the subjects were high school students. Each subject was given twelve sets of tests. The tests consisted of a five-place, a six-place, a seven-place and an eight-place number span. Three sets of tests were presented orally at one sitting to the subjects. The digits were administered at the rate of two numbers per five seconds of time. After the experimenter read each number, the subject was required to give immediate written reproduction. It was concluded that the memory span increases with age rather than with the growth of intelligence, as determined by the tests used in promoting pupils from one grade to another.[2]

The answer to the question "How large a quantity of a given sort of material can be reproduced perfectly after one presentation?" was answered by Jacobs (1887). He devised a method using digit span series from three digits to twelve digits. Jacobs instructions are similar to those used in the Stanford-Binet -

> "I will say some numbers; when I have finished, you are to repeat the numbers in the same order."[3]

---

[2]T. L. Bolton, "Growth of Memory in School Children," American Journal of Psychology, (1871), IV, 362-380.

[3]J. Jacobs, American Journal of Psychology, (1928), LV, pp. 285-290.

The score in this experiment is the length of the list which the subject can recall and recite perfectly after only one presentation.

Jacobs, in his research, found that 1) individual methods of grouping are used in memory span tests for the longer series; 2) attention span is a term which satisfactorily describes the immediate reproduction of separate descrete units; 3) a brief inspection indicates little difference in test results whether digits are presented in haphazard or serial order; 4) attention span, for most individuals, is short enough to cause grasping of units after the number five has been passed; 5) the attention span ranges from 2.5 to 8, while the memory span ranges from 6 to 3.5.

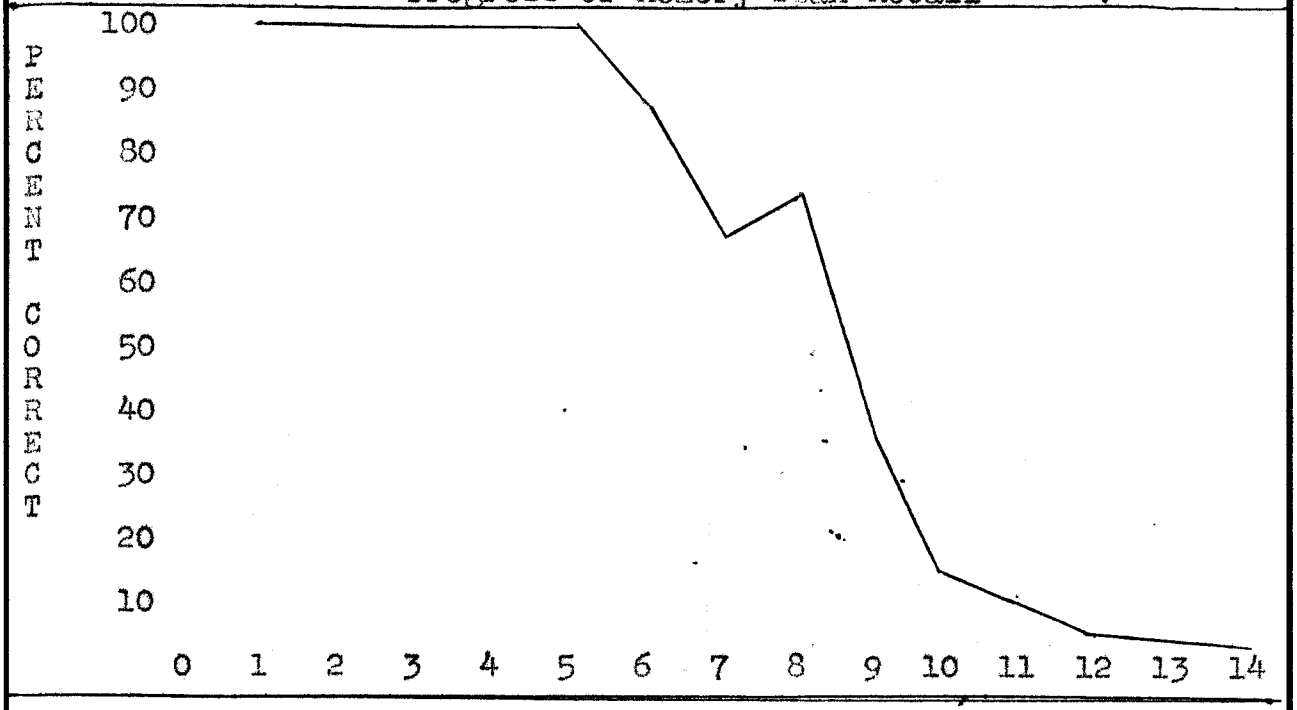
Oberly's study brings out the fact that the individual memory span is not a fixed quantity, but varies from trial to trial. The following graph shows that the subjects of Oberly's experiment recited perfectly memory span lists of two digits, three digits, four digits, and five digits, but after a span of five digits, the percent of those able to give perfect responses declined gradually as the memory span lists were lengthened.[4]

---

[4]H. S. Oberly, American Journal of Psychology, (1928), XXXV, pp. 295-302.

TABLE II

Progress of Memory Span Recall



An experiment was performed by Woodworth and Poffenberger, using lists of ten digits to determine percentage of errors in memory (serial) learning. The lists were read aloud by the experimenter, after which the group of twenty-six subjects wrote their recollections of each list presented. The fewest errors occurred in the first serial position.[5]

Testing auditory memory is also a major part of the Revised Stanford-Binet, as was it in the original Binet Scale.

---

[5]R. S. Woodworth and A. T. Poffenberger, Textbook of Experimental Psychology, Mineog. ed., (New York, 1920), pp. 71-72.

According to Terman, the auditory memory for digits test do not measure pure memory because it is impossible to isolate any function for separate testing. In any test, general ability is operative.

In the Stanford-Binet, a digit span test first appears at the two year-six month level. This is a span of two digits. Three digits are attained by the average person at the three year level. Four digits are repeated successfully by the average person of four years-six months and five digits by the average seven year old. The addition of one number to the five digit span test on the Binet is equal to three years of mental development.

The digit span test has been retained in the Binet in spite of its tendency to arouse negativism in some children. It has a very good increase in the percent of children passing from one age to the next and its correlation with the composite score is .62.

The method of administration in the Stanford-Binet is for the examiner to say, "I am going to say some numbers, and when I am through, I want you to say them just the way I do. Listen carefully and get them just right." The digits should be pronounced distinctly and with perfectly uniform emphasis, at the rate of one per second.[6]

---

[6]Lewis M. Terman, The Measurement of Intelligence, Haughton Mifflin Company, (Cambridge, Mass. 1944), pp. 194-199.

Unlike the Stanford-Binet method, Wechsler, in his intelligence tests, has grouped the digit span tests. This was done for two reasons. One was because of the limited range of them when taken separately. On memory span for digits forward, a score range of only four points includes about ninety percent of the adult population, and about the same percentage for digits backwards. By combining the scores obtainable on both into one test measure, it was possible to extend the test's range, as well as to close wide gaps obtained between successive scores when tests are used singly.

The second reason for grouping the tests was for the purpose of reducing the amount which each contributed to the total score. It was found that the Digit Span Test with a total score has a reliability of .51. Wechsler concluded that a good rote memory is of practical value but correlates very little with higher levels of intelligence.

In Wechsler's tests, the digit span tests begin with a span of three digits and increases to nine digits. The subjects are given two attempts at success and after two successive failures on one span, the testing ceases on this item. Wechsler stated that he desired at one time to eliminate the digit span tests from his tests, but for the following reasons, they were retained.

1. While memory span for digits is, on the whole, a poor measure of intelligence, it is nevertheless a good one at the

lower levels. Except in cases of special defects or organic diseased adults who can retain only five digits forward and three backwards will be found, in nine cases out of ten, to be feebleminded.

2. Special difficulty with the repetition of digits forward or backward is often of diagnostic significance.[7]

Klugman's study is restricted to the establishing of the best placement of the Digits Test in the Wechsler-Bellvue Scale.

These are the hypotheses to be investigated: 1) The digits test should be "given first when the applicant is relatively free from fatigue." 2) The best placement may be in the middle of the battery following a warming-up period before the effects of fatigue are felt. 3) It may be that the end position in the battery is the best one because of the "better adjustment of the testee to the testing situation" as the examination proceeds.

Three hundred white native-born psycho-neurotic returnees referred by medical officers for W-B examination were tested. The men were tested in random order and no effort was made to control the order of the other tests in the battery.

In order to determine whether the digits scores were

---

[7]David Wechsler, The Measurement of Adult Intelligence, The Williams and Wilkins Company, (Baltimore, 1944), pp. 83-85.

affected by position in the scale, three method variations were employed. The digits were administered at the beginning, in the middle, and at the end of the battery.

There are several practical reasons why an answer to this problem would be desirable. First, if method differences exist, a subject would not be penalized by getting a lower IQ score. Second, diagnostic and clinical procedure would give more certain and meaningful data. Third, calculation of deterioration would yield more accurate scores. Fourth, the digits test could be included or omitted from a shorter form of the Wechsler-Bellvue with greater validity.

Klugman drew the following conclusions from his investigation.

1. The mean digit results of Forward 6.05, Backward 4.38, and Total 10.43, are in agreement with those found by another investigator.

2. The order of desirability of position in the scale appears to be Middle, Beginning, and End.

3. The Digits Backward were not affected differently from Digits Forward.

4. To what extent the obtained results are due to the uncontrolled factor of what tests preceded and followed the administration of the Digits sub-test is not known.

5. Whether these results are true only for psychoneurotic individuals under the given circumstances or for the

population as a whole, remains for future research. Similarly, whether the same type of findings exist for other sub-tests shall depend on results obtained from other studies.[8]

R. L. Newton did a comparative study of two methods of administration of the Digit Span Test. A comparison of two methods of administration of the digit span test of the Wechsler-Bellvue Intelligence Scale, Form I, was made using twenty-eight hospitalized patients at the V. A. Hospital in Aspinwal, Pennaylvania. The following results were obtained.

1. The pitch of the voice on the last digit of a series significantly affected the results obtained in the subject's score. If the pitch of the voice was lowered, the subject recalled more digits accurately.

2. The pitch of the voice significantly affected the digits forward, but did not have a significant effect upon the digits reversed.[9]

Stanley Maldowsky and Patricia Corcoran Maldowsky conducted an experiment to decide whether the digit span was an anxiety indicator. The experiment was designed to test the hypothesis that anxiety will function so as to cause a significantly greater decrement in Digit Span scores than in Vocabulary Scores.

---

[8]Samuel F. Klugman, Journal of Consulting Psychology, (1940), XII, pp. 345-348.

[9]R. L. Newton, Journal of Clinical Psychology, (1950), VI, pp. 409-412.

Thirty-two college students acted as subjects. These students previously tested on the Full Scale Wechsler-Bellvue were retested with the Vocabulary and Digit Span sub-tests. One-half of the subjects received the usual clinical rapport-establishing procedure. These subjects were called Control Groups I and II. The other half of the subjects received the anxiety-inducing procedure. These subjects were labeled Experimental Groups I and II.

Control Group I and Experimental Group I received Digit Span first and Vocabulary second. Control Group II and Experimental Group II received the opposite order of presentation. The groups were matched according to Verbal I. Q.

The results supported this hypothesis. They would tend to reinforce the clinicians confidence in the Digit Span sub-test as being sensitive to situational anxiety and in the Vocabulary sub-test as being relatively impervious to it.[10]

The means of the weighted scores for both sub-tests are presented in Table III.

Table IV represents the mean differences in the tests. The mean results of the Control Groups were combined, as were the results of the Experimental Groups, in Table IV.

---

[10] Stanley Maldowsky and Patricia Corcoran Maldowsky, Journal of Consulting Psychology, (1952), XVI, pp. 115-118.

TABLE III[11]

Test-Retest Differences of the Four Sub-Groups Studied by Maldowsky

	Test I	σ	Retest	σ	Difference
Control Group I					
Vocabulary	12.62	1.17	13.00	1.00	+.38
Digit Span	11.75	2.16	12.50	2.64	+.75
Control Group II					
Vocabulary	12.50	.50	13.25	.83	+ .75
Digit Span	12.13	2.17	13.13	2.28	+1.00
Experimental Group I					
Vocabulary	11.63	.85	13.38	.66	+.75
Digit Span	12.13	1.96	12.00	2.96	-.13
Experimental Group II					
Vocabulary	12.13	.78	12.63	.68	+ .50
Digit Span	12.88	2.14	10.75	2.77	-2.13

TABLE IV[12]

Mean Test - Re-test Differences Of Group Studied by Maldowsky

Test	Controls Mean	σ	Experimentals Mean	σ
Vocabulary	+.562	.622	+.437	.747
Digit Span	+.375	8.234	-1.125	8.859

[11] Ibid., p. 116.

[12] Ibid.

Robert G. Bernreuter, in an unpublished article on the "Improvement of Memory Tests," gave the results of the experimentation being done with his new memory span test. He used the digit span test, the letter span test, and the nonsense syllable span test. A total of 190 fourth grade children were tested. The method of presentation and scoring was similar to that used on the Stanford-Binet. There were two forms to the tests and each was given to each child in twelve different sequences in order to control the effect of any possible difference in difficulty between the forms and to control the effect of practice and fatigue. An analysis of variance was then made which showed that the effect of practice and fatigue was negligible. It also showed that the forms were well equalled.[13]

The results show that for the number span test the mean length was 5.83, and for letters 4.89 and for nonsense syllables 3.71.

The extent to which the scaling of the test could be refined can be shown by ratios of the scoring unit to the standard deviation. When numbers, Form A were administered, a standard deviation of .94 was obtained, but when both Forms of numbers were administered to the subjects, a standard

---

[13]Robert G. Bernreuter, "Improvement of Memory Span Tests," APA Meeting, (September, 1953) Unpublished article, p. 3.

deviation of .53 was obtained. Letters, Form A, had 1.12 standard deviation and when Form A and B were combined, .63 was obtained. Syllables, Form A had a 1.47 standard deviation and a .89 for Form A plus B. Combining Numbers and Letters gave a .31 standard deviation while the combining of the three tests gave a .25 standard deviation.

The lengthening of the test not only improved its discriminatory value, it also increased the reliability. This can be shown by the reliability coefficients. The reliability of Numbers Form A vs. Form B was .69, but when they were combined, the reliability was raised to .82. The reliability of Letters alone was .51, but when combined, .68. A simple Syllables test had a reliability of .44, but the two forms together had a reliability of .61. When both forms of numbers and letters were combined, a reliability coefficient of .87 was obtained. The addition of syllables did not further raise the reliability.

Bernreuter concluded that a memory span test can be constructed using two forms of a number test and two forms of a letters test, which would be reliable and sufficiently discriminatory to be used in individual diagnosis.[14]

In the latest study on the Bernreuter Test, forty children in kindergarten and nursery school were given the test. Twenty-six children enrolled in a kindergarten, ranging in age

---

[14] Ibid., p. 4.

from 5-0 to 6-0 years, and fourteen children in nursery school, from 3-8 to 5-0 were used. The mean age for the groups was approximately 5-5 and 4-6, respectively. In spite of the difference of almost a year, the ranges differed by only one point, and the difference between the means of the two groups was only .2. For this reason, the scores were combined. It is suspected that the groups were highly selected in regard to intelligence. Twelve of the subjects had been given the Stanford-Binet, and their scores varied from 115 to 140.

The kindergarten children were tested at two separate times; the nursery children all in one session. The order used was N, L, S-A, N, L, S-B. The average time of testing ran about ten minutes each, with about half the time spent in the actual testing, and the rest in rapport building procedures. Interruptions to discuss dogs, cowboy boots, etc. were frequent.

Six additional children were seen but not tested due to negativism or interruptions. The vast majority seemed to enjoy the "game" and tried to do well. There seems to be evidence of fatigue or monotony on Form B, which at this age level seems to offset any practice effect. Of the forty children, twenty-four scored one or two points higher on Form A, six one point higher on B, and ten the same. There was a tendency for the children to begin to repeat the longer series before the examiner had finished.

The scores range from 16 to 31, and are fairly normally

distributed. The N range varies from 3 to 5, L from 2 to 6, and S from 1 to 5. The mean for N is slightly higher than 4, for L slightly less than 4, and almost 3 for S.

It was possible to secure twelve Stanford-Binet IQ's and so to compare M. A.'s and the test scores. The validity of the M. A.'s is not guaranteed since most of them were practice tests. However, there seems to be a rough correlation.[15]

---

[15]Joan Fagan, Report on the Bernreuter Memory Span Test, Spring, 1955. Unpublished study at Pennsylvania State College made under direction of Robert Bernreuter.

CHAPTER III

THE PROCEDURE

In order to investigate the reliability of Bernreuter's Memory Span Test the following procedure was used.

THE SUBJECTS. The subjects of this experiment were 114 children in the sixth grade in parochial schools of Chicago, ranging in age from eleven to thirteen years. Forty-eight were boys and sixty-six were girls. The mean age was 11.4 years and a standard deviation was .49. Each child was tested individually by the experimenter. Each child received the Bernreuter Memory Span Test.

THE MATERIALS. The materials that were to be learned consisted of two distinct and separate forms of the Bernreuter Memory Span Test, Form A and Form B.

A pilot study was undertaken by Robert G. Bernreuter and Pennsylvania State College in 1953. At this time new testing materials were developed and an attempt was made to determine the reliability and fineness of scaling of these new materials.

In the first study, letters and nonsense syllables were substituted for digits. It was found that confusion was created when letters ending in an "ee" sound were used. A sequence such as "b, p, d, c, g," would often be repeated as "b, p, e, d, c," even by children whose memory span was considerably longer than five units. As a consequence all such

letters were omitted. The letters "f, h, k, l, n, r, s, and q" were finally selected.

The first syllables tried were the typical three letter nonsense syllables consisting of two consonants and an intervening vowel, which were similar to those constructed by Ebbinghaus. This type of syllable was found to be useful in memory tests when presented visually, but they were found to be very confusing when presented orally. Instead, nonsense syllables were constructed by adding an "aw" sound to the letters b, d, f, g, h, k, m, and t. The result was that the following syllables were formed; "baw, daw, faw, gaw, haw, maw, and taw."

On the basis of the first study conducted by Bernreuter, six new memory span tests were constructed. The six new memory span tests were divided into two groups. These groups were classified as Form A and Form B. Each form consisted of a number span test, a letter span test, and a nonsense syllable span test. The difference between Form A and Form B does not rest on a difference in sequence of presentation, but simply a difference in the spans themselves. For example, in Form A, in the number span test, the first series of three digits is 385, and the first series of three digits in Form B begins with 528.

Care was taken by Bernreuter in the construction of the test to avoid starting a series with the same symbol with

which the immediately preceding series had started or ended. This means that all series which might follow immediately after the series "F S Q H" could not start with either F or H. Care was also used to avoid repeating pairs of letters in the series immediately following. Thus, after "F S Q H," the combinations FS, SQ, and QH could not be used. These precautions were found necessary, according to Bernreuter, to keep the series approximately equal in difficulty.

In administering the Bernreuter Memory Span Test, the child was first given the number span test of Form A. The experimenter began with one digit span and as soon as the child passed a span of one length, the experimenter moved to the next longer span. Thus, if he correctly repeated 385, the next span given was 2947. If the child failed 385, but passed the next span, 197, the experimenter would go on to 2947. If the child failed 385 and failed 197, but passed 624, the experimenter went on to 2947. However, if the child failed 385, 197 and 624, the experimenter stopped the administration of the number span test and went on to the letter span test. This procedure was used for both Form A and Form B, and for each span test - numbers, letters and nonsense syllables.

The method of presentation and scoring was similar to that used on the Stanford-Binet. Success was based on the

correct repetition of one of three trials.

The Memory Span Tests were administered in the following order -

Number Span - Form A
Letter Span - Form A
Syllable Span - Form A

As soon as the three sequences were completed, the experimenter administered Form B.

Number Span - Form B
Letter Span - Form B
Syllable Span - Form B

Both forms of each test were administered to each child in the above sequences. The purpose of administering the sequences in this order was to control the effect of any possible differences in difficulty between the forms and to control the effect of practice and fatigue.

In the following experiment, numbers will appear as N, letters as L, syllables as S, and when the forms are combined, the results will appear as T; eg. NT would stand for Numbers, Form A and Form B.

The following instructions were given to each subject before the test began:

> "This is to see how well you can listen. I am going to say something. I want you to listen and, when I am through, I want you to tell me what I said. We will start with numbers."

If there was no response, the experimenter said "You tell me what I said."

The subject was encouraged to try, if necessary. Only the single letters, numbers, and syllables could be repeated; the longer series could not. The letters, numbers, or syllables were read at the rate of one per second. When the subject tried and passed, the first series on the next higher level was given. A success was considered if only one of the series of three was passed. If the subject passed the first series of the total series, the examiner immediately went to the next higher series. When the subject tried, and failed, the next series of the same level was given. This was continued until there had been three consecutive failures in a level.

Each subject was examined in a small room in the school. The room was set away from the school classrooms. Each child was seated in an arm chair with a table in front of him. The examiner was seated directly across from the subject.

The lighting was florescent. The noise was at a minimum. In order to control fatigue, each child was tested in the morning. It was impossible to control the tone of the examiner's voice when each test was administered, but he spoke in normal conversation pitch. It was also impossible to control the possibility of a hearing defect or loss in the subjects.

CHAPTER IV

ANALYSIS OF RESULTS AND DISCUSSION

Each subject's successes and failures were examined carefully and the total score for each subject was obtained for each test of both forms. These results will be discussed in relation to a) the frequency responses on the different spans; b) the differences in means for the various tests; c) the ranges of spans; d) the extent to which the test can be refined; e) its discriminatory value and reliability; f) whether the three types of tasks utilize three kinds of memory.

In Appendix I, the frequency of responses for the letters, numbers, and syllables span tests are shown. The results represented by this appendix indicate the highest successes obtained by subjects before three consecutive failures.

It should be noted that in Appendix I, the scores for both forms clustered around the five and six letter spans. On Form B there was a slight improvement in length of span repeated over Form A. Eleven boys repeated the six letter span test perfectly in Form A, whereas fourteen repeated it perfectly in Form B. The same percentage increase happened for the seven letter span test. This increase could possibly be due to practice and familiarity with the letters. But before any definite conclusion could be made, further investigation would be needed to determine the reason for the increase from Form A to Form B.

From the results in Appendix I, it should be noted that performance changed very little. The only notable change was in the increase of subjects basing at the four letter span length in Form B.

The boys' results on the numbers test indicate little or no improvement from Form A to Form B. For four subjects there was regression to a lower span level. For one subject, there was an increase in span length.

The girls' results on the numbers test are rather interesting. In Form A, the results were more scattered from the mean, and span length was higher. But in Form B, the subjects' scores clustered more around the mean of 6.46. In Form A, four subjects obtained scores of nine and fourteen scores of five, but in Form B, six of the subjects that obtained a score of five moved closer to the mean as did the four individuals that scored nine on Form A.

The syllable span tests were the only ones in which there was a marked regression in span length. The boy that scored seven on Form A received a score of five on Form B.

The regression for the girls on the syllables span test was not quite so noticeable as the boys, but there was still regression. The girls tended to cluster more closely around the mean than the boys. But the deviation between the two groups was minute.

The mean memory span length for the 114 subjects on each

test are shown in Table V. As can be seen, the mean length of memory span when numbers were used was approximately six for both forms, separately and combined. When letters were used, the mean length was approximately five; and, when syllables were used, the mean length was approximately four.

The mean length of memory span for numbers and letters combined was approximately six, but when syllables were added the span length dropped to approximately five.

Without further investigation, it would be inadvisable to reverse the test items in the administration of the Bernreuter Test. There would perhaps be need to investigate other factors which might be operating when a shift in presentation of materials occurs; orientation which involves presenting what appears to be more difficult items before presenting those which are apparently simple may have bearing on the success or, inversely, lack of success with the test items. Therefore, it would be unwise to draw any conclusion or defend any position concerning the reversal of test items without further investigation.

The ranges of span for numbers are 4-9, letters 3-8, and syllables 2-7. This indicates that numbers are easier to repeat than letters or syllables, and that letters are easier to repeat than syllables.

TABLE V

Means and Standard Deviation for Numbers, Letters, and Nonsense Syllables

Test Span	Means Form A	σ Form A	Means Form A + B	σ Form A + B
N	6.4	1.17	6.38	1.17
L	5.41	.93	5.41	.93
S	3.91	.98	3.84	.88
N + L	5.9		5.87	1.15
N + L + S	5.24		5.19	1.38

The above table presents the standard deviations of the distributions of Form A and Form A plus Form B. Table V shows that a change of one digit in the Form A numbers equals 1.17 standard deviations. This, of course, is very coarse scaling. When the forms are combined, the standard deviation did not change at all. The same held true for letters when combined. As shown, a change of one letter in Form A equalled .93 standard deviation with no change in standard deviation when tests were combined. On the Form A Syllables, a change of one syllable equals .93 standard deviation. However, the situation changed slightly when the syllable tests were combined, showing that the change of one syllable equalled .88 standard deviation.

The above results indicated that Form A and B of the number tests were comparable. The same could be said of the letters and syllables.

When the number and letter tests were combined, the standard deviation was 1.15, indicating that the group was not as homeogenous as Bernreuter's group. When he combined the two tests, the standard deviation was .31,[1] indicating a rather homeogenous group. This held true when the three tests were combined.

The lengthening of the test increased the reliability of the test. This is shown in Table VI.

Reliability Coefficients of Parallel Forms

Test Span	Form A vs. Form B	Form A plus Form B
N	.58	.73
L	.47	.63
S	.41	.58
N + L	.71	.83
N + L + S	.69	.83

The reliability of a single form of the numbers test was found to be .58. This is about what is generally reported in the literature. The digit span test in the Wechsler-Bellvue Intelligence Scale for Adults has a reliability of .51.

When Form A and Form B of the number span test were combined, the reliability for Parallel Forms using the

---

[1]Bernreuter, op. cit., p. 3.

Spearman-Brown Formula was increased to .73. When letters were added to numbers, the reliability for Parallel Forms was raised to .83. The adding of syllables to the other two tests did not further increase the reliability. It decreased slightly.

The fact that series of letters are easier to repeat than are series of syllables, and that series of numbers are still easier to repeat, raises the question of whether the three types of tests utilize three different kinds of memory. Inter-correlations give only a partial answer to this question.

The correlations vary from .63 to .79. When letters, Form A and B were correlated with syllables, Form A and B, the correlation obtained was .63. This was somewhat different than the results obtained in a previous study. The lowest correlation obtained was between numbers and syllables. The correlation between numbers and letters was .79, indicating that letters and numbers, besides increasing the reliability of the test, probably are measuring a common element. From the results it can be stated that it is uncertain whether the syllables test involves some additional factors.

When the new correlations were corrected for attenuation, the results obtained gave further indication that numbers and letters measure a common element and that syllables do not.

The above results indicate this clearly. When Numbers, Form A and B were combined to Letters, Form A and B, the

correlation was .79. When these new correlations were corrected for attenuation, the reliability was increased to .95. Whenever numbers were combined to syllables and letters combined to syllables, the reliability was raised to .74 and .75 respectively. This indicates that these forms are not as reliable as the combining of numbers and letters and do not measure a common element.

CHAPTER V

SUMMARY AND CONCLUSION

The only type of memory span test in which there has been much research is the number span test. Binet reported that children of ten years should have a span of six digits. The results obtained confirmed this, since a mean of 6.38 was obtained. The average for the 114 subjects was 11.4 years, which would indicate a memory span of slightly above six digits; this was arrived at by using the Terman-Merrill method of standardization on the Stanford-Binet. In their standardization, the repetition of six digits is characteristic of the ten year old child.

Little or no research has been done concerning letter or syllable span tests. From indications, it seems that individuals of eleven years repeat one less letter than number and two less nonsense syllables than numbers.

Looking at the span ranges, it would indicate that numbers are easier to repeat than letters, and letters easier than syllables. But to state that an individual has a mean memory span length of 6.38, 5.41, or 3.84 and a range length from four to nine, is meaningless and useless unless reference is made to the type of material utilized.

Adding syllables to letters and numbers seems somewhat pointless since there is no increase in the reliability to any significant extent. For this reason, it would seem that to

delete the syllable test in the construction of new memory span material would be advisable.

One may wonder why an analysis of variance was not undertaken to determine what effect practice and fatigue had on the subjects. This was undertaken in three previous studies with the Bernreuter Memory Span Test and the results indicate that the effect of practice and fatigue were negligible. This could be understandable since the test is administered in under ten minutes and the subjects never repeat the same material.

The final analysis of the test results and introspections suggest three specific changes in the letter span tests.

The first change is the elimination of the letter "K." Forty percent of the subjects, at one time or other, repeated this letter as "A." When this was first noticed, the experimenter took extra precautions in enunciating this letter. It made little difference since the letter "K" was still repeated as "A."

Another improvement in the letter span tests is the elimination of "Q." It seems that the youngsters tested had been drilled in their reading classes that "Q" never is alone. "U" always follows it. This error was repeated by twenty-seven percent of the subjects. For example, in this span, L N S F Q K, it would be repeated as L N S F Q U K.

Another frequent error repeated by many subjects was the

substitution of "M" for "N." This is due to the marked similarity in their sounds.

If a letter span test is to be formed with these eliminations, new letters are to be found. An example of such a span might be F, H, L, R, S, T, and W. This would eliminate letters ending in an "E" sound. The only exception to this would be "T," but this letter is rarely misunderstood. It would also eliminate vowels and letters sounding similar to other letters.

This experiment was conducted in order to investigate methods of improving memory span tests.

Three tests containing two forms of a number span test, a letter span test, and a nonsense syllable test were presented once to 114 sixth grade subjects individually. The recall of the items was required to be made in the original order. All errors were recorded on the prepared blanks.

The general conclusions are as follows:

1. Memory span length varies with the type of material used.

2. It was found feasible to construct new tests substituting letters for numbers. The attempt to use nonsense syllables seems not to be very satisfactory.

3. It also could be concluded that a memory span test could be constructed using two forms of a number test and two forms of a letter test. These would be sufficiently reliable

and sufficiently discriminatory to be used in individual diagnosis of memory span.

4. It becomes apparent that some corrections would be advisable in the letter span test. Elimination of "K, Q, and M" is one possibility.

5. It seems that further work is needed to determine whether the syllables test involves some additional factor not yet recognized.

6. It becomes evident that the interpretation by clinicians based on a single form of a digits forward memory span test is not justified.

7. Finally, it is apparent that further investigation concerning a shift in method of presentation is advisable. This is needed in order to discover what factors are involved when supposedly more difficult materials (syllables) are presented first and simple materials following in presentation.

BIBLIOGRAPHY

Bernreuter, Robert G. "Improvement of Memory Span Tests," Unpublished article given at the Pennsylvania State College, (September, 1953).

Bolton, T. L. American Journal of Psychology. IV, 1891.

Ebbinghaus, Herman. Memory. Trans. by Henry Ruzer and Clara E. Bussenus, New York, 1913.

Fagan, Joan. Unpublished report on the Bernreuter Memory Span Test, (Spring, 1955).

Guilford, J. P. Fundamental Statistics in Psychology and Education. New York, 1950.

Guthrie, E. R. The Psychology of Learning. 1935.

Hull, C. L. American Journal of Psychology. VL, 1935.

Jacobs, J. American Journal of Psychology. VL, 1922.

Klugman, Samuel F. Journal of Consulting Psychology. XII, 1940.

Lindquist, E. F. A First Course in Statistics. Boston, 1938.

Maldowsky, Stanley and Maldowsky, Patricia Corcoran. Journal of Consulting Psychology. XVI, 1952.

McGeoch, John A. and Irion, Arthur L. The Psychology of Human Learning. New York, 1952.

Newton, R. L. Journal of Clinical Psychology. VI, 1950.

Oberly, H. S. American Journal of Psychology. XXXV, 1928.

Sargent, S. Stanfeld. Basic Teachings of the Great Psychologists. New York, 1954.

Terman, Lewis M. The Measurement of Intelligence. Cambridge, Mass., 1944.

Wechsler, David. The Measurement of Adult Intelligence. Baltimore, 1944.

Woodworth, Robert S. Contemporary Schools of Psychology. New York, 1948.

Woodworth, R. S. and Poffenberger, A. T. Textbook of Experimental Psychology. Mimeog. ed., 1920.

Woodworth, Robert S. and Schlosberg. Experimental Psychology. Revised ed., New York, 1954.

APPENDIX I

Frequency of Span Length on the Letter, Number, and Syllable Span Tests, Form A and Form B Obtained by Boys and Girls

	SPAN LEVEL							
	2	3	4	5	6	7	8	9
Letters, Form A								
Boys	0	0	9	23	11	4	1	0
Girls	0	0	6	32	18	8	2	0
Letters, Form B								
Boys	0	1	6	20	14	7	0	0
Girls	0	0	8	33	16	8	1	0
Numbers, Form A								
Boys	0	0	3	11	18	8	7	1
Girls	0	0	1	14	22	14	11	4
Numbers, Form B								
Boys	0	0	6	9	16	7	7	2
Girls	0	0	1	8	26	21	10	0
Syllables, Form A								
Boys	0	13	29	5	0	1	0	0
Girls	1	13	42	10	0	0	0	0
Syllables, Form B								
Boys	0	20	23	5	0	0	0	0
Girls	0	20	38	7	1	0	0	0

APPENDIX II

Bernreuter Memory Span Test
Letters, Form A

R
L
F

QS
HN
KR

NKH
QLF
HRN

KNRL
FSQH
RNKS

RQNHF
NLHQS
QKNHF

HSQFRL
QLRHNS
RHFKNL

NQLSRFK
KRLFNSH
SLHFRNK

QNFLRKHS
KHQNRLSF
RFSHNKQL

NHKQLRFSQ
RLFQSKHNF
KFHLRNSQL

LHFRNSQKRH
NRFSHLKQFR
QNKRSFHLSK

LA
LB

NA
NB

SA
SB

LT
NT
ST

AT
BT
T

LANAT
LBNBT
LNT

Name . . . . . .

School . . . . . .

Grade . . . . . .

Birthdate . . . . . .

Date . . . . . .

Age . . . . . .

Sex . . . . . .

Order of presentation . . . .

Examiner . . . . . .

Score . . . . . .

Remarks:

APPENDIX III

Bernreuter Memory Span Test
Numbers, Form A

4	LA	Name . . . . .
3	LB	
1		School . . . . .
	NA	
8 6	NB	Grade . . . . .
7 2		
5 9	SA	Birthdate . . . . .
	SB	
3 8 5		Date . . . . .
1 9 7	LT	
6 2 4	NT	Age . . . . .
	ST	
2 9 4 7		Sex . . . . .
3 1 5 8	AT	
4 6 1 3	BT	Order of presentation . . . .
	T	
9 2 5 7 4		Examiner . . . . .
1 4 6 3 8	LANAT	
7 9 5 2 6	LBNBT	Score . . . . .
	LNT	
4 6 1 8 5 2		Remarks:
9 3 8 1 4 2		
5 8 2 4 7 3		
5 7 1 3 6 2 4		
6 1 7 4 2 9 5		
8 5 4 9 2 6 1		
9 1 3 5 2 7 4 8		
7 4 1 5 3 8 2 9		
3 8 6 2 4 7 1 5		
2 5 7 1 9 6 4 8 3		
6 1 4 9 3 7 2 5 8		
1 8 3 6 2 4 9 5 7		

APPENDIX IV

Bernreuter Memory Span Test
Syllables, Form A

HAW
BAW
KAW

GAW FAW
DAW MAW
TAW BAW

HAW BAW FAW
DAW KAW TAW
MAW GAW BAW

KAW BAW TAW GAW
FAW MAW HAW DAW
TAW GAW MAW HAW

HAW GAW TAW FAW KAW
DAW TAW MAW GAW BAW
TAW FAW KAW GAW DAW

KAW TAW DAW BAW MAW GAW
DAW GAW FAW TAW KAW BAW
HAW FAW KAW TAW DAW MAW

BAW DAW MAW HAW FAW GAW KAW
DAW MAW HAW BAW FAW TAW GAW
TAW DAW FAW BAW HAW GAW MAW

GAW KAW TAW BAW MAW DAW HAW FAW
KAW HAW GAW FAW MAW TAW DAW BAW
FAW BAW MAW GAW HAW KAW DAW TAW

HAW FAW DAW MAW TAW KAW GAW BAW DAW
BAW GAW KAW FAW TAW HAW DAW MAW FAW
KAW BAW TAW HAW MAW FAW DAW GAW TAW

GAW BAW KAW FAW TAW DAW HAW MAW TAW KAW
HAW TAW GAW HAW BAW FAW MAW DAW BAW GAW
FAW GAW MAW TAW KAW DAW BAW HAW TAW MAW

LA
LB

NA
NB

SA
SB

LT
NT
ST

AT
BT
T

LANAT
LBNBT
LNT

Name . . . . . .

School . . . . . .

Grade . . . . . .

Birthdate . . . . . .

Date . . . . . .

Age . . . . . .

Sex . . . . . .

Order of presentation . . . .

Examiner . . . . . .

Score . . . . . .

Remarks:

APPENDIX V

Bernreuter Memory Span Test
Letters, Form B

S
H
R

NK
QF
LS

FSQ
KQS
NHL

LQKR
FNSH
QRLK

NLHFS
RLNHQ
HRNQF

LNSFQK
SLFKHR
NHQFSK

QFKNHLS
LKFQHRN
SLRQNFK

NHLKRFQS
QKHFSRNL
RNFHSLQK

KNLFHSRQL
SRNKHLQFK
HKFRNSLQF

RKNSHQFLNK
SNHKRFQLRH
QFHSKNRLSF

LA
LB

NA
NB

SA
SB

LT
NT
ST

AT
BT
T

LANAT
LBNBT
LNT

Name . . . . . .

School . . . . . .

Grade . . . . . .

Birthdate . . . . . .

Date . . . . . .

Age . . . . . .

Sex . . . . . .

Order of presentation . . . .

Examiner . . . . . .

Score . . . . . .

Remarks:

APPENDIX VI

Bernreuter Memory Span Test Numbers, Form B

8	LA	Name . . . . .
2	LB	
4		School . . . . .
	NA	
1 5	NB	Grade . . . . .
7 3		
6 2	SA	Birthdate . . . . .
	SB	
5 2 8		Date . . . . .
6 3 9	LT	
2 7 4	NT	Age . . . . .
	ST	
9 5 3 6		Sex . . . . .
8 1 4 7	AT	
3 5 1 9	BT	Order of presentation . . . .
	T	
5 2 4 9 7		Examiner . . . . .
6 1 5 2 4	LANAT	
1 7 2 6 9	LBNBT	Score . . . . .
	LNT	
4 3 9 1 6 5		Remarks:
7 5 1 6 2 9		
3 6 4 8 2 7		
8 1 6 9 3 5 7		
5 3 1 8 2 6 4		
1 9 6 8 3 7 2		
6 2 7 1 4 9 5 3		
3 7 5 2 4 8 6 1		
9 3 1 5 2 6 4 7		
2 5 1 6 4 8 3 7 9		
3 1 6 9 2 5 7 4 8		
4 8 5 3 7 1 6 9 2		

APPENDIX VII

Bernreuter Memory Span Test
Syllables, Form B

TAW
MAW
KAW

BAW FAW
DAW BAW
HAW MAW

GAW BAW HAW
FAW KAW DAW
MAW HAW TAW

KAW DAW GAW BAW
FAW TAW BAW HAW
BAW DAW GAW MAW

TAW FAW KAW BAW GAW
KAW HAW TAW GAW MAW
BAW MAW KAW DAW TAW

HAW BAW GAW MAW FAW KAW
DAW MAW BAW HAW GAW FAW
MAW HAW DAW BAW KAW TAW

BAW DAW MAW KAW FAW GAW HAW
KAW GAW FAW HAW BAW MAW TAW
GAW MAW DAW KAW HAW TAW FAW

HAW GAW TAW FAW BAW KAW DAW MAW
TAW FAW KAW HAW GAW DAW MAW BAW
GAW BAW HAW DAW KAW TAW MAW FAW

DAW FAW GAW TAW KAW BAW MAW HAW TAW
HAW GAW DAW FAW KAW TAW BAW MAW DAW
FAW KAW TAW HAW BAW MAW GAW DAW KAW

BAW HAW GAW MAW KAW DAW TAW FAW KAW GAW
FAW GAW KAW TAW DAW MAW HAW BAW TAW KAW
HAW MAW FAW KAW GAW TAW BAW DAW FAW MAW

LA
LB

NA
NB

SA
SB

LT
NT
ST

AT
BT
T

LANAT
LBNBT
LNT

Name . . . . . .
School . . . . . .
Grade . . . . . .
Birthdate . . . . . .
Date . . . . . .
Age . . . . . .
Sex . . . . . .
Order of presentation . .
Examiner . . . . . .
Score . . . . . .
Remarks:

APPROVAL SHEET

The thesis submitted by Joseph Michael Angileri has been read and approved by three members of the Department of Psychology.

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval with reference to content, form, and mechanical accuracy.

The thesis is therefore accepted in partial fulfillment of the requirements for the Degree of Master of Arts.

June 5, 1957
Date

Frank J. Kobler
Signature of Adviser