VARIABILITY IN MEMORY SPAN*

DAVID MITCHELL

New York City

THE memory span of individuals of the same age has been known to vary in quite marked degree. A person of very low grade intelligence may have a memory span of 3, 2, or 1 or may even be unable to reproduce one impression. On the other hand exceptional individuals have been able to reproduce 10, 11, 12 or even a greater number of discrete impressions. In the same way they vary in every known function. Measurements of ability to read show that in the same age group there are wide variations. These variations are such that in the ten year group there are children who are not able to read any better than some children in the nine year group. Some children of the eight year group excel some of the twelve year group. In the different age groups there is overlapping, the children of one age varying so much that always some of the one age are better than some of those, one, two, three, or four years older.

In a large measure we have tended to ignore this variability and to think that the average performance was typical of every child. The child of eight years should be able to spell certain words, should be able to do certain mathematical operations and should be able to read paragraphs containing words of more or less well known difficulty. All children have been supposed to reach and none have been supposed to greatly exceed this average performance ability. Only recently has this status of fixed averages for age groups been seriously questioned.

The variability in an individual's performance has to a much greater extent been overlooked. Increase in performance ability, corresponding with age increase, has been recognized, but that a child should be better in a performance one day than he would be in a similar performance on the following day has not been generally conceded. If a child is able to read a certain passage on Monday, he should be able to read it as well on Tuesday, even if he has had no practice in the interval, and should read it better if an opportunity for practice has been provided. Furthermore, if he has been able to spell a certain group of words the first day of the month, we

*Paper read at The New York Branch of the American Psychological Association, November, 1918.

have assumed that he should be able to spell those words with the same degree of accuracy at the end of the month. If he has been able to repeat correctly the multiplication table on Friday, his ability to do this on the following Monday has not been questioned. Our norms of performance ability have been established on the basis of a belief in the constancy of the individual's performance.

The constancy of average results has confirmed us in our notion of the individual's constancy. The repeated testing of groups of individuals has given average results and measures of deviation which were approximately identical from day to day and from group to group. A study of the individual child and his status from day to day in the group might have shaken our faith in the reliability of individual results, but so far little study of this kind has been made. In order to evaluate the assumption of the constancy of an individual's performance, it was decided to test one function which is supposed to be directly dependent upon neurone excitability. This function cannot be materially changed by practice although it does improve with age. The reproduction of discrete elements, being considered a function of the organic memory, should show as little variation as a function such as mathematical calculation or spelling a certain number of words, since performance in mathematical calculation and in spelling depends largely upon the training or practice which one has had.

Subjects

The children chosen for the experiment were included in five fifth grade groups in one school and three seventh grade groups in other schools. In the five fifth grade classes there were over two hundred children. In the three seventh grade classes there were over one hundred and twenty-five children. Of the fifth grade children, 192 were given a Memory Span test for digits on two days which were separated by a period of exactly five weeks. Of the seventh grade children, 97 were given the Memory Span test on ten successive school days. The fifth grade group included 103 girls and 89 boys; the seventh grade classes 53 boys and 44 girls.

Procedure

The series of digits for each of the trials were made up by a chance selection of numbers, care being taken to eliminate sequences either forward or backward. The presentation of the numbers was given

orally to the group in the various classes and the reproductions were written by the pupils. As the experimenter entered the room at the time of a regular class session, the teacher distributed the record blanks and asked the children to clear their desks and to be certain to have pencils ready. The experimenter then gave instructions in the following way: "We are going to do an experiment this morning. In an experiment we try to find out something. I am going to try to find out how many numbers you can remember. In an experiment it is important that everyone should do the same thing in the same way, so I am going to ask you to pay very close attention to what I tell you and then do exactly as I say. First I want you to write your name at the top of this sheet of paper which has been given to you. Then write your age, your grade, the date, and whether you are a boy or girl." When the children had completed these instructions the experimenter proceeded. "I am going to say some numbers to you and after I have finished saying them I want you to write them down. For instance, if I should say 2-4-6, you would write down at the top of the paper 246 (indicating the method by writing the numbers on the board). Then if I should say 5-7-9, you would write 579 under the other number in this way (indicating them on the board in the right position). You would do this for as many numbers as I might give you. Do you understand? It is important that no one should start to write until I have given all the numbers, so I am going to ask you to hold up your pencils like this (the experimenter holds his pencil up at the level of his shoulder). I will also hold my pencil up and when I bring it down like this (indicating by the movement of the hand to the level of the desk), that will be a signal to you to write down the numbers I have just given you. As soon as you have finished writing them down I want you to hold up your pencil again and that will be a signal to me that you are ready for the next series of numbers. Do you all understand? Then let's begin. Everyone put his pencil up. First I will give you three numbers. Ready!" The test begun, the succeeding series of numbers were given as soon as all the pencils had been raised. The numbers were dictated as distinctly as possible in a monotonous tone of voice at the rate of one per second. The series given were two trials on each series of from 3 to 10 digits.

A method of scoring the Memory Span was a matter of serious question. Consideration was given to the method of scoring suggested by the modified Spearman's Foot rule Correlation $r = 1 - \frac{\sum d^2}{(n^2-1)/3}$.

The amount of labor involved prohibited the use of this method. We therefore, adopted an arbitrary ruling. Two scores were made. The first score was the longest series, in both attempts at which the child was successful at reproduction and in no series previous to which had he failed. This score we may call the "Consecutive Series" score. The other or "Maximum Series" score is the longest series which the child reproduced, regardless of the number on which he had previously failed. A third method of scoring might have been adopted which would give as the child's score the longest series on which the child correctly reproduced one of the trials, and on no series previous to which had he failed on both the trials. Since the question was the variability and not the determination of an absolute standard for different groups the two scores mentioned seemed sufficient.

Results

Tables I and II present the "Consecutive Series" score and the "Maximum Series" score respectively for the 192 children who were tested on two days separated by the period of five weeks.

TABLE I
"Consecutive Series" Score

Trial	II	3	4	5	6	7	8	9	10	Total
I										
3		2	2	1						5
4		5	16	23	5					49
5		5	20	26	13					64
6		2	6	21	21	10	1		1	62
7				2	6	3				11
8			1							1
9										
10										
Total	14	45	73	45	13	1	1			192

¹WHIPPLE, G. M. *Manual of Mental and Physical Tests*. Vol. II, p. 161.

TABLE II
"Maximum Series" Score

Trial	II	3	4	5	6	7	8	9	10	Total
I										
3										
4			<u>2</u>	<u>2</u>	<u>2</u>					6
5			<u>1</u>	<u>13</u>	<u>18</u>	<u>3</u>		<u>1</u>		36
6			<u>2</u>	<u>17</u>	<u>40</u>	<u>18</u>	<u>8</u>	<u>4</u>	<u>1</u>	90
7				<u>2</u>	<u>7</u>	<u>9</u>	<u>4</u>	<u>4</u>		26
8					<u>1</u>	<u>6</u>	<u>3</u>	<u>1</u>		11
9				<u>1</u>	<u>2</u>		<u>1</u>	<u>1</u>	<u>2</u>	7
10			<u>1</u>	<u>2</u>	<u>3</u>	<u>3</u>	<u>1</u>	<u>2</u>	<u>4</u>	16
Total			6	37	73	39	17	13	7	192

These are correlation tables showing the number of children who made the various scores on the two trials. The scores for trial I are shown in the first column. The scores for trial II are shown in the first line of the table, beginning with the second column. In the table of "Consecutive Series" scores we find that two children had a score of three on each of the two trials; two children scored three on the first trial and four on the second trial; one child whose score was three on the first trial scored five on the second trial. In the third line and third column of Table II we find that two children had a score of four in the "Maximum Series" on the first trial and four on the second trial. The next column of the same line shows that two children scored four and five on the first and second trials respectively. Two children also scored four and six in the first and second trial respectively. In both tables the number of children who had the same score on the two trials is underlined.

The number of children who gave the same score on the two trials, or whose score differed by one, two, or three or more numbers is shown in Table III.

TABLE III
Relation of Scores on Two Trials

"Consecutive Series"		"Maximum Series"	
Score		Score	
68	Equal Score	72	
100	Different by One	79	
20	Different by Two	23	
2	Different by Three	9	
2	Different by Four	6	
0	Different by Five	2	
0	Different by Six	1	

The first column is for the "Consecutive Series" score. It shows that 68 children had the same score on the two trials. One hundred children scored one more in one trial than they did in the other. Twenty children had scores on the two trials different by two while four children's scores differed by three and four. Similar results are presented in the final column for the "Maximum Series" score. By this method of scoring the variation in the two trials is greater. There were only 151 children whose score was within one of being the same for the two trials. Of the other children 18 varied three or more, two of them varying as much as five and one as much as six.

Table IV shows the record of all children with the averages and medians for both methods of procedure and each trial. The first column of the table gives the scores. The second and third columns contain the numbers of children making the various "Consecutive Series" scores on the two trials.

TABLE IV
Number of Children Making Different Scores

Score	"Consecutive Series" Score		"Maximum Series" Score	
	Trial		Trial	
	1st	2nd	1st	2nd
3	5	14	0	0
4	49	45	6	6
5	64	73	36	37
6	62	45	90	73
7	11	13	26	39
8	1	1	11	17
9	0	0	7	13
10	0	1	16	7
Average	5.1	5.0	6.4	6.5
Median	5	5	6	6

Columns four and five contain the corresponding "Maximum Series" score. The numbers in columns two and three are not identical nor are the numbers in columns four and five. They are sufficiently alike to make the average and the median almost the same for the two trials. The median for the "Consecutive Series" score is the same in both trials and the average differs by only 0.1.

In the "Maximum Series" score the median is again the same, the average score in the two trials differing by only 0.1. The average or group performance for the two days is practically identical. We are not, however, justified in saying that the children show a constancy of performance since they have been shown in Table III to vary considerably.

Tables V and VI show results for the 97 children who were tested on ten successive days.

TABLE V.
"Consecutive Series" Score.

		Actual Scores on the Different Trials										0	Total
Average Scores and Number of Children Making Them.		3	4	5	6	7	8	9	10				
3.0													
3.5	1	1	6	1	1							1	10
4.0	4	4	24	7	4							1	40
4.5	10	5	36	37	21							1	100
5.0	23	5	52	77	82	12	1					1	230
5.5	26	2	32	48	127	41	8					2	260
6.0	16		13	14	80	40	13						160
6.5	5		4	2	15	19	9	1					50
7.0	9		1	3	20	29	30	7					90
7.5	3		2		3	9	7	6	3				30
8.0													
8.5													
9.0													
9.5													
10.0													
Total		17	170	189	353	150	68	14	3	6			970

In the first part of the first column, the average scores in nearest half units are shown:—that is, if a child made a score of four on six of the ten days, a score of three on two other days, a score of five on the remaining two days, his average score for the ten days would be four. Four children, whose average score individually was four, had scores similar to these. In the second part of the column, the number of children who made the different average scores is shown. In the first line of the table the different scores are shown. For example, in the line for the average score of 7.5, it is shown that three children made this average. Two of the 30 trials gave scores of four, three of six, nine of seven, seven of eight, six of nine, and

three of the trials gave scores of ten. The last line in the column gives the number of times each of the different scores was made in the 970 trials made by these 97 children.

The construction of Table VI is identical, this table being for the "Maximum Series" score.

TABLE VI.
"Maximum Series" Score
Actual Score on the Different Trials

Average Score and Number of Children Making Them.	3	4	5	6	7	8	9	10	0 Total
3.0									
3.5									
4.0									
4.5									
5.0	2	3	8	8	1				20
5.5	1	1	2	7					10
6.0	12	2	14	61	37	6			120
6.5	20		11	84	71	26	7	1	200
7.0	17		2	38	74	43	11	2	170
7.5	17			17	47	81	23	2	170
8.0	9			5	13	37	27	8	90
8.5	9				6	30	37	17	90
9.0	7					15	31	24	70
9.5	3					1	10	19	30
10.0									
Total		6	37	220	249	239	146	73	970

As the average "Maximum Series" score increases, we see that the numbers are sharply cut off in the last column of the table. It was assumed in the beginning of the experiment that very few children would correctly reproduce a series of ten digits. We find that 73 out of the 970 trials result in a "Maximum Series" score of 10. It is obvious from these results that the series should have been increased in length to at least twelve, and possibly a higher number.

The question is raised as to the average daily score in these ten days. Table VII presents the result.

TABLE VII.
Average Daily Scores

Day	"Consecutive Series"	"Maximum Series"
	Score	Score
1	5.5	6.9
2	5.8	7.2
3	4.8	7.3
4	5.6	7.2
5	5.8	7.4
6	6.0	7.6
7	6.0	7.8
8	5.7	7.6
9	5.9	7.6
10	6.0	7.8

In the first column of this table the day on which the trial was given is shown. The second column gives the average "Consecutive Series" score for the days, and the third column gives the average "Maximum Series" score. In the first case the range of scores is from 4.8 to 6.0, a range of 1.2. This range is wide and should be kept in mind when considering the later discussion of correlations in which the results of trial III are correlated with results from each of the other trials. The reason for this lower score which is 0.8 lower than any other score in the column is found in the fact that the experimenter for one group of children changed his method of procedure for that day. He consciously lengthened the time between the enunciation of the digits in a series. The result was quite marked;—not one of the children with whom he worked succeeded in getting a "Consecutive Series" score higher than four on that day.

The range of variability in averages with the "Maximum Series" score is somewhat less. The lowest score is 6.9 and the highest is 7.8, a range of 0.9. The group results are similar from day to day as was the case with the 192 children whose averages for the "Consecutive Series" score differed by only 0.1 and the averages of the "Maximum Series" score showed only the same difference. Group results are approximately identical from day to day but individual scores vary considerably.

In order to get a simpler indicator of individual variabilities coefficients of correlation were calculated according to the formula

$$r = \frac{p^2}{\sigma \times \sigma y}$$

¹See page 181ff. *Introduction to the Theory of Statistics*, by G. UDNY YULF.

TABLE VIII.
Coefficients of Correlation

Trials	"Consecutive Series"	"Maximum Series"
	Scores	Scores
I-II	0.40	0.54
I-III	0.09	0.49
I-IV	0.46	0.50
I-V	0.30	0.59
I-VI	0.19	0.64
I-VII	0.31	0.59
I-VIII	0.31	0.50
I-IX	0.38	0.50
I-X	0.46	0.52
II-III	0.30	0.49
II-IV	0.45	0.61
II-V	0.47	0.71
II-VI	0.42	0.55
II-VII	0.45	0.62
II-VIII	0.27	0.49
II-IX	0.42	0.48
II-X	0.60	0.53
III-IV	0.41	0.51
III-V	0.44	0.57
III-VI	0.46	0.61
III-VII	0.18	0.63
III-VIII	0.17	0.46
III-IX	0.26	0.58
III-X	0.16	0.52
IV-V	0.60	0.60
IV-VI	0.49	0.58
IV-VII	0.37	0.62
IV-VIII	0.41	0.44
IV-IX	0.47	0.67
IV-X	0.38	0.46
V-VI	0.33	0.65
V-VII	0.47	0.73
V-VIII	0.38	0.57
V-IX	0.53	0.62
V-X	0.51	0.51
VI-VII	0.28	0.67
VI-VIII	0.33	0.63
VI-IX	0.39	0.58
VI-X	0.43	0.65
VII-VIII	0.29	0.56
VII-IX	0.44	0.63
VII-X	0.54	0.49
VIII-IX	0.48	0.40
VIII-X	0.35	0.57
IX-X	0.64	0.48
Average	0.37	0.54

The correlations for the two trials with the 192 children were 0.47 and 0.44 for the two methods of scoring. For the ten trials with the group of 97 children we have a total of 45 coefficients of correlation. They are given in order in Table VIII.

The first column gives the trials correlated and the second and third columns the coefficients for "Consecutive Series" and "Maximum Series" scores respectively. The values are somewhat less than they should be for one reason and somewhat more than they should be for another. As mentioned before, one of the three experimenters changed his method for the third trial so that not one of forty children scored higher than four on the "Consecutive Series." All these correlations, therefore, in which the third trial is considered are lower than we might expect. On the other hand, our series of digits did not go lower than three. A number of children failed in getting this series correct so that their "Consecutive Series" score could only be zero. Since this break occurred in the series it did not seem wise to include the zero score in the correlation. The coefficients are, therefore, somewhat higher than we might expect if series of digits down to one had been used.

A collation of the coefficients is shown in Table IX.

TABLE IX
Coefficients of Correlation

Coefficients of Correlation	"Consecutive Series"	"Maximum Series"
0.00-0.09	1	
0.10-0.19	4	
0.20-0.29	4	
0.30-0.39	12	
0.40-0.49	18	10
0.50-0.59	3	18
0.60-0.69	3	15
0.70-0.79		2
Total	45	45

The first column shows them grouped in ranges of ten points. The second and third columns contain the number of the correlations. According to this table the coefficient of correlation for one pair of trials was less than 0.9. For four trials it ranged between 0.10 and 0.19; for another four between 0.20 and 0.29. In only six cases

is the coefficient of correlation for "Consecutive Series" scores greater than 0.50. The average is 0.37. The average coefficient of correlation for the "Maximum Series" score is 0.54. Thirty-three come between 0.50 and 0.69; only two are greater than 0.70. If the memory span were a constant function, one would be justified in expecting a higher correlation than those that are shown. A coefficient less than 0.75 is extremely significant when different functions are being correlated, but correlations of less than that value of the same function on different days leads us to suspect that the apparent validity of the individual test may be seriously questioned.

Conclusion

In our evaluation of memory span we have been largely concerned with group or average results. We have determined what the score for children in a great many different problems should be. We have assumed, if a child reaches that average score in one examination, that he has the ability to do it on all occasions. We, therefore, claim that he ranks in intelligence in a certain part of the intelligence scale. On the other hand if in one examination a child fails to make the average score, we have judged him incapable of doing it and ranked him in a lower part of the intelligence scale. Our results indicate that this method of grading needs revision. The performance ability of children is not a static or fixed quantity. It rather varies from month to month, from day to day, or we might even say from hour to hour. A child who is able to make 100% in Spelling one day may make only 75% or 50% in Spelling the same list of words on another day. Children who are able to do 35, 40 or 50 problems a minute one day may do considerably less or considerably more on another day.

When we consider the question from the standpoint of average performance for a group, we do not need to be concerned with this variability. The average score for a group remains approximately constant with the same degree of practice, or increases in fairly constant ratios with similar amounts of practice. Children in one city with the same amount of practice as children in another city give almost identical results. But the individual child may materially change his own position in the rank order. He who is first may not change his position so that he is last but he may change it to such an extent that neither the individual score nor the average

of two scores will give a satisfactory record of his ability or his intelligence, or even of his capability in the specific function tested.

These results may not be considered as conclusive evidence that our mental testing is being done under the handicap of a false assumption. It indicates however, that a revision of our ideas concerning it may be necessary. In the scales of intelligence which have been devised, great weight is placed upon the ability to reproduce numbers and also to reproduce series of words in sentences. The method of giving the test and of scoring probably eliminates some of the variability which we have found, but when the "Maximum Series" score of children can vary as much as six within the range of ten, it does seem that injustice may be done to the child who on a certain day gives his lowest possible score in the memory span. Before the evidence may be considered conclusive, it will be necessary to determine the amount of variability when the method followed is identical with that prescribed in the intelligence scale. There is little doubt, however, that the variability in performance merits careful and serious consideration and demonstrates that the doubt should always be in favor of the child until a wider range of experiment with him has been possible.