

Journal of Experimental Psychology

VOL. II, No. 5.

OCTOBER, 1917

A REPETITION OF EBERT AND MEUMANN'S PRACTICE EXPERIMENT ON MEMORY¹

BY H. B. REED

University of Idaho

Reviewers of Ebert and Meumann's practice experiment upon memory are generally agreed that the experiment must be repeated before their conclusion can be accepted as valid for the facts of the experiment. Lack of space prevents me from taking notice of the reviews and criticism that have been made on their work. The references of Müller, Sleight, and Thorndike cited below will be found valuable in this connection.

Suffice it to say that the general conclusion drawn from the results was that a special training of the memory improves it in general. But since the amount of transfer from the special training to the tests was not equal in all the tests, a particular conclusion was drawn in that the amount of transfer follows the law "that the special memories are improved exactly to the degree that the learning materials and learning methods are related to the specially trained memory" (p. 200). These conclusions have not been accepted and have been criticized among others by G. E. Müller² and R. Wessely³ in Germany, by W. G. Sleight⁴ in England, and W. F. Dearborn⁵ and E. L. Thorndike⁶ in America.

So far as is known to the writer, Meumann took no notice of his critics except G. E. Müller whom he ridicules for his

¹ *Archiv. f. ges. Psychol.*, 4, 1904, 1-232.

² *Zeit. f. Psychol.*, 39, 1905, 111-125.

³ *Neues Jahrb. f. Päd.*, 16, 1905, 296-309; 371-386.

⁴ *British Jour. Psych.*, 4, 1911, 386-456.

⁵ *Psych. Bull.*, 6, 1909, 44.

⁶ *Psych. of Learn.*, 1913, 369-376.

overemphasis upon "precision." He states that if his own explanation of transfer is mysterious, that of Müller¹ is equally so. In confirmation of his conclusions, he refers to the experiments of Winch,² Daniel Starch,³ Coover and Angell,⁴ and Wallace Wallin,⁵ which he reviews as the 'Pädagogisch wichtigste.' Nowhere does he refer to the experiment of Thorndike and Woodworth⁶ which was published three years before his own, although the remarks upon special memories in the introduction of the report of his experiment have a striking similarity to the opening remarks of Thorndike and Woodworth upon the same subject. Nor does he take notice of other experiments reaching different conclusions, *e. g.*, those of James,⁷ Fracker,⁸ Klein,⁹ Foster,¹⁰ Ruger,¹¹ and Sleight,¹² all of which are reviewed in a book to which he makes more than one favorable reference.¹³ It appears, therefore, that Meumann never retracted the fundamental conclusions of his experiment.

Since this experiment of Meumann forms the background of most that is said in his 'Psychology of Learning,' and of much that is said in his 'Vorlesungen,' and has elsewhere occupied a prominent place in psychological literature, and coming as it does from the hand of one of the foremost psychologists of his time, it appeared to the writer that it ought to be repeated, and that in this way many of the important criticisms that have been made upon it might be answered. I shall refer to this repetition as the test experiment which I now wish to describe.

There were two series of tests: I., memory-span tests, and, II., learning tests.

¹ *Vorlesungen*, III., 1914, p. 143 ff.

² *British J. of Psych.*, 2, 1908, 284-293; 3, 1910, 386-405.

³ *J. of Ed. Psych.*, 3, 1912, 209-213.

⁴ *Amer. J. of Psych.*, 18, 1907, 237-340.

⁵ *J. of Ed. Psych. Mon.*, 1911.

⁶ *Psych. Rev.*, 8, 1901, 247-261; 384-395.

⁷ *Principles of Psych.*, I, 1893, 667.

⁸ *Psych. Rev. Mon. Suppl.*, 4, 1908, 56-102.

⁹ Klein, Thorndike, 'Psych. of Learn.,' 401 ff.

¹⁰ *Jour. of Ed. Psych.*, 2, 1911, 11-21.

¹¹ *Archives of Psych.*, No. 15, 1910.

¹² *British Jour. of Psych.*, 4, 1911, 386-456.

¹³ Thorndike's 'Psych. of Learn.'

I. The memory-span tests consisted of:

1. Thirteen double series of consonants ranging from five to seventeen in the first test. In the second, the quantity was increased to twenty letters.

2. Sixteen double series of numbers ranging from five to twenty. In the second test, the largest quantity was seventeen.

3. Eight double series of nonsense syllables ranging from four to eleven.

4. Twelve series of disconnected one-syllable words ranging from four to fifteen.

5. Nine series of Latin-English vocabularies ranging from three to eleven pairs, tested by the Treffer method.

6. Passages from Wordsworth's 'She was a phantom of Delight,' 'The 'Solitary Reaper,' 'Stepping Westward,' and 'To a Highland Girl,' with the following number of words per passage in the first test: 12, 14, 16, 18, 21, 22, 24, 26 and 28. In the second test, the quantities were: 11, 14, 16, 18, 20, 22, 24, 26, and 28.

7. Eleven passages of prose from Locke's 'Essay,' ranging from twelve to thirty-four words.

The first four of the above tests were presented in an auditory form, at the rate of one member a second. The last three tests were also read aloud at a slow rate, but not at the rate of one member per second. The reproduction in all cases was written upon paper by the subject immediately after the stimulus ceased.

II. The learning tests were all presented visually and were relearned after twenty-four hours. They consisted of:

8. Three series of nonsense syllables, consisting of ten, twelve, and sixteen syllables respectively. The syllables were made from the English alphabet, and according to the conditions of Müller and Schumann.

9. Two series of meaningless visual diagrams of twelve members each, the same as those used by Ebert and Meumann.

10. Thirty Latin-English vocabularies.

11. Forty Latin-English vocabularies.

12. Two eight-line stanzas of poetry selected from Wordsworth's 'To the Daisy.'

13. A ten-line passage from Locke's 'Essay.'

The experimenter's function in the learning test was to see that the subject carried out his instructions, and to keep account of the number of repetitions made and of the time, which was taken with a stop watch. The time was measured from the beginning of the test to the end of the first correct reproduction. The same measurements were also kept for the syllable rows in the practice series. In the calculation of the data, only the times were taken into account.

The test experiment was made with fourteen adult subjects, but later reduced to thirteen, at the University of Wyoming during the summer school session of 1915. Eight of these constituted Group I., and the other six, later reduced to five, constituted Group II.

The eight subjects of Group I. did both the practice and the test series of the experiment while the six of Group II., later reduced to five, omitted the practice series but did the test series at the same time as Group I. Group II. was the control group. The two groups were divided with respect to equal learning ability as accurately as this could be determined by the experimenter's personal judgment, but, as a matter of fact, the average learning ability of Group II. was slightly superior, which makes the results for control purposes all the better.

Like the experiment of Ebert and Meumann, the test experiment had a practice series which consisted of learning and relearning twelve-syllable nonsense series, constructed as explained above. Those which were learned one day were relearned twenty-four hours later. The form of presentation was somewhat different from that in Test 8. Instead of the experimenter allowing the subject to reproduce the syllables at his own rate, he kept presenting them in a mechanical way at the rate of one per second until the learner succeeded in reproducing for the first time each syllable before it was turned. The series was then considered learned. The four methods of learning: the whole, the part, and the two mixed methods, followed by Ebert and Meumann's subjects, were also followed by the subjects of Group I. of the test experiment. The order

of the different learning methods was varied irregularly in the test experiment. Ebert and Meumann kept the quantity of learning material constant for each practice exercise so that each subject learned two series of syllables daily. But in the test experiment, the constancy was put upon the amount of time. This was limited to thirty minutes for each exercise, except when the subject failed to learn one twelve-syllable row. In that interval, each subject relearned what he had learned the day before, and then learned as much new material as possible in the space of time remaining. Four of the subjects in Group I. had two practice exercises daily, one in the morning and one in the afternoon, while the other four had only one practice exercise daily in the morning. Several of the subjects omitted Sundays and other holidays. Approximately fifteen days were spent in practice by Group I., while upon the five days immediately before and after the practice, Groups I. and II. took the tests described above. The test experiment may therefore be considered as repeating only the first practice period of Ebert and Meumann together with the tests that came before and after that period. If this part of the experiment is confirmed or refuted, the writer is of the opinion that a repetition of the remaining parts would not essentially change the character of the results.

The course of improvement in the training series is represented in the practice curves of Fig. 1 and Fig. 2. The points in the ordinates are based upon the average learning and relearning times per series for each day of practice. That the character of the improvement in learning nonsense syllables, which is memorial or ideational in function, is of the same sort as that of learning telegraphy or typewriting and of other sensori-motor functions is evident from the character of the curves, which are typical learning curves. The individual differences, however, are striking.

The relearning curves either vary about a horizontal line or else run closely parallel to the learning curves. Contrary to our expectation from the results of Ebbinghaus, Müller, and Schumann, Radosawljewitch on the relation of the amount saved to the strength of the imprinting, a high learning time

is in most cases not followed by a correspondingly low learning time, but rather the converse. But it should be said that no special light was thrown upon the causes of improvement. In the cases under observation, they appeared not to consist in lack of effort or interest, or the development of bad learning methods such as unusual pronunciation. Nor did the coöperation of the will appear to be of absolute and controlling importance, as it is stated by Meumann, for no exertion of the will prevented bad records or plateaus. See his 'Psychology of Learning,' pp. 64, 66, 74, 115, 139, 283, 287, and 361.

But whatever may be the cause of improvement, the curves of Fig. 1 and Fig. 2 show that a very great improvement was made by the subjects of the test experiment in learning nonsense syllables. The question now is: How much of this improvement was transferred to other mental functions?

The measures employed by Ebert and Meumann to determine the effect of the improvement gained in learning nonsense syllables upon the memory-span functions were the memory-span, the first excess of one third error, and the first excess of one half error. The measure of the memory-span was taken as the number of members in the largest series correctly reproduced in any given test. The amount of error was calculated as follows: "The omission or addition of a member was considered as $\frac{4}{4}$ error, a displacement in the series of more than one position as $\frac{3}{4}$ error, and the displacement of the same sort of exactly one position as $\frac{2}{4}$ error and a correction as $\frac{1}{4}$ error (*op. cit.*, p. 11)."

In the test experiment, the three above measures were used in the same way. But to increase the reliability of the tests five additional measures were used as follows:

1. The *method of average error* per quantity on series in each test calculated according to the scale of Ebert and Meumann for each series. From the results thus obtained, the average was calculated for all the series in a test together with the average deviation from the individual series.
2. The *method of retained members*, according to which the number of members correctly reproduced regardless of order and position was calculated in each series of a test. From

these results the average and the average deviation were calculated as in I.

3. The *method of retained groups*, according to which the largest number of members reproduced that had the same

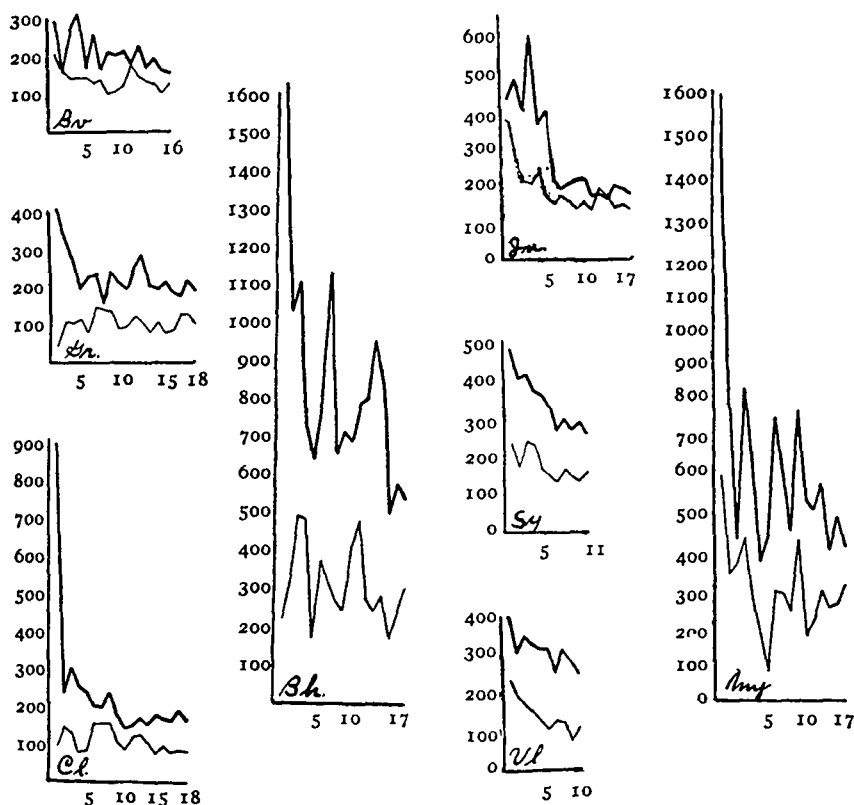


FIG. 1.

FIG. 2.

FIGS. 1 AND 2. Practice Curves Showing Course of Improvement in Learning and Relearning Twelve-syllable Rows of Nonsense Syllables. The time in seconds is plotted along the ordinates and successive practice days along the abscissas. The heavy line above represents the average learning time per series from day to day; and the light line below, the corresponding relearning time.

order of succession as those in the original series was calculated for each series in a test. The average and average deviation were determined as in I.

4. The *method of retained positions*, according to which the number of members in the reproduction that had the

same position relative to either the first or the last end of the series was calculated. The average without the A.D. was determined as in 1.

5. The method of correlation by the Pearson Coefficient, the application of which will be described later.

In Table I., the results of the above measures except the last one are given from the various tests for Group 1 and for Group 2.

The first column in Table I., 'Highest No. Correct,' represents the memory-span measurements. The second, third, fourth, sixth, eighth, and tenth contain the measurements in the order described above. Test I. represents the tests made before the practice series and Test II. represents the tests made after the practice series. The averages opposite I. and II. are the averages, and the A.D.'s under them are the deviations, of the individual averages from the group averages. D_1 represents the differences between tests I. and II. for Group 1, while D_2 represents the corresponding differences for Group 2. A plus indicates the amount of gain in Test II. over Test I. A minus indicates a corresponding loss. $D_1 - D_2$ the difference in gain or loss between groups 1 and 2. The A.D.'s in the vertical column after the rubrics, 'Error per Quantity,' 'Mems. Repr. per Quantity,' etc., are the deviations of the individual subjects in the test in question. The reason that the A.D.'s of 'The Error per Quantity' are so large is that the increasing the quantity of learning material beyond the memory-span increases the amount of error enormously. Thus, for example, when five letters were read the amount of error in the reproduction for most of the subjects was zero, but when a series of twenty letters were read, the amount of error for most of them was between sixteen and seventeen. This feature will become clearer when we shall discuss the influence of the quantity of learning material upon the reproduction.

From Table I. it will be seen that the value of the 28 gains of Group 1 in Test II. is destroyed by the counterbalance of 19 losses, by the fact that only two of them were more than twice as great as the A.D.'s of the averages from which they

TABLE I—Continued

WORDS—Concluded

Group	Test		Highest No. Correct	First Excess of $\frac{1}{2}$ Error	First Excess of $\frac{1}{2}$ Error	Error per Quan- tity	Aver. A.D.	Mems. Repr. per Quan- tity	Aver. A.D.	Size Largest Group Repr. per Quan- tity	Aver. A.D.	No. of Posi- tions Correc- tly Repr. per Quan- tity
2	I	Ave.	5.67	7.17	7.33	6.57	4.11	5.42	.91	2.93	1.01	3.69
		A.D.	.56	.89	1.11	.72	.31	.41	.27	.46	.13	.57
	II	Ave.	3.17	6.50	6.50	6.95	4.04	5.00	.80	2.53	1.00	3.27
		A.D.	.28	.67	.83	.45	.56	.33	.26	.42	.09	.58
		D_1-D_2	-.50	-.67	-.83	-.38		-.42		-.40		-.42
			+.37	+.42	+.33	-.53		+.67		+.35		+.79

VOCABULARIES

1	I	Ave.	4.64	4.71	4.71	4.35	2.38	3.27	1.02	3.45	.97	3.18
		A.D.	.23	.08	.08	.52	.10	1.16	.21	.96	.48	1.04
	II	Ave.	4.29	4.30	5.14	3.67	1.96	3.87	1.10	2.60	.99	3.57
		A.D.	1.40	1.11	1.63	.97	.29	1.00	.21	.94	.43	1.44
		D_1	-.35	-.41	+.43	+.68		+.60		+.15		+.39
2	I	Ave.	3.80	5.00	5.20	4.17	2.38	3.05	.78	2.23	.86	2.42
		A.D.	1.00	1.20	1.04	1.49	.36	.54	.31	.52	.08	.77
	II	Ave.	3.40	4.80	5.60	4.06	2.38	3.58	.74	2.31	.82	2.89
		A.D.	.88	.64	1.12	.68	.44	.40	.11	.60	.11	.80
		D_1	-.40	-.20	+.40	-.09		+.53		+.08		+.47
		D_1-D_2	+.75	+.21	+.03	+.77		+.07		+.07		-.08

POETRY

1	I	Ave.	18.55	18.28	19.13	9.21	7.87	14.29	3.24	9.88	4.08	11.76
		A.D.	2.19	1.07	1.62	3.34	1.47	1.71	.67	2.06	1.11	2.68
	II	Ave.	10.32	14.86	16.30	12.72	6.43	13.11	2.93	8.02	2.60	10.35
		A.D.	4.40	2.45	3.19	3.07	1.02	1.99	.85	2.36	.86	5.73
		D_1	-8.23	-3.42	-2.83	-3.51		-.18		-1.86		-1.41
2	I	Ave.	18.60	20.80	20.80	7.66	7.01	11.75	3.38	16.00	3.10	14.22
		A.D.	2.28	2.24	2.24	1.60	1.15	1.65	.64	1.35	.40	1.83
	II	Ave.	14.20	17.20	17.60	12.25	8.17	14.06	3.52	9.35	2.84	11.33
		A.D.	1.52	1.44	1.92	2.37	1.86	1.36	.46	1.31	.32	2.21
		D_1	-4.40	-3.40	-3.20	-4.59		+2.31		-6.65		-2.89
		D_1-D_2	-4.83	+.18	+.37	+1.08		-2.49		+4.79		+1.48

PROSE

1	I	Ave.	15.00	17.00	17.50	15.04	9.97	14.59	8.44	8.84	3.46	10.36
		A.D.	1.50	1.50	1.50	1.44	.66	1.55	.35	1.13	.86	1.52
	II	Ave.	17.00	20.00	20.50	13.07	8.09	16.43	3.86	10.70	3.50	12.92
		A.D.	3.50	2.50	2.15	2.15	.79	1.55	.34	1.99	.82	2.06
		D_1	+2.00	+3.00	+3.00	1.97		1.84		1.86		2.56
2	I	Ave.	16.66	16.33	21.33	14.70	1.24	14.90	3.08	9.40	3.48	10.48
		A.D.	1.88	2.33	3.33	2.56	.74	3.12	.47	1.75	.18	1.46
	II	Ave.	20.00	22.00	23.20	12.45	8.66	17.23	3.71	11.05	3.54	11.55
		A.D.	3.20	.00	1.92	3.23	1.52	2.05	.95	2.23	.54	1.68
		D_1	+3.34	+5.27	+1.87	+1.25		2.33		1.65		1.07
		D_1-D_2	-1.84	-2.67	+1.13	+.72		+.49		+.21		+.49

were computed, and by the 19 gains made by Group 2 in the same tests. When the difference in gain and loss between the groups is considered, we discover that Group 1 had thirty-nine gains over Group 2, and against these ten losses must be counted. This situation in itself apparently indicates that Group 1 increased its memory-span somewhat by the training in nonsense syllables, but, when the character of the gains and losses which make up this situation are considered, the facts of Table II. can only mean that the *special training had no effect in either increasing or decreasing the memory-span* for the various materials with which the tests were made. The important condition of this conclusion is that neither Group 1 nor Group 2 made any significant gains or losses in Test II. as compared with Test I.

Table II. shows the average performances of Groups 1 and 2 in the learning tests. *L* = Learning; *RL* = Relearning; ' = minutes and '' = seconds. The other symbols are the same as in Table I. The A.D.'s represent the deviation of the individuals of the group from the group average.

Table II. gives the measurements of the average performances of Groups 1 and 2 in the learning tests. Group 1 made three gains and three losses in the learning functions. Only one of the gains is greater than the A.D. of the average of either Test I. or Test II., and the other two are less. Group 2 made three gains and three losses in the learning functions, and all of them are less than the A.D.'s of the averages of either Test I. or Test II. This can only mean that there is clearly no *general learning ability that is improved by the special training in nonsense syllables*. Although no general learning ability is improved by the special training in nonsense syllables, it is still possible that this training might have improved some special learning function. In the test with nonsense syllables, Group 1 gained 6' 47.3'' over Group 2. In the diagrams, Group 1 gained 5' 12.9'' over Group 2. These gains are much greater than those gained in any other functions. When these gains are not considered with reference to the A.D.'s, they indicate that the special training did improve the ability to learn nonsense syllables and nonsense

visual diagrams. We must bear in mind, however, that Group 1 might gain more in a given test than Group 2, without the greater amount of improvement being due to special training.

In the relearning functions, Group 1 has five gains and one loss, but in all cases, at least one of the A.D.'s of the average from which these gains or losses were computed, is greater than the gain or loss in question. Group 2 has four gains and two losses. Only one of the gains is greater than the A.D. of the group averages from which it was computed. These facts undoubtedly mean not only that there is no general relearning ability which is improved by special training in learning nonsense syllables, but also that the efficiency of no special learning function is apparently affected by the special training.

The measurements given in Tables I. and II. may be supplemented by the measurement of the Pearson coefficient of correlation. If an individual's memory-span and his learning ability are allied, or if both are allied to some general ability, and if one of these factors is improved by a special training series, there ought to be a positive correlation between the performances in the training series and the performances in the test series, and the correlation between the training series and the second test series ought to be higher than that between the training series and the first test series.

Since there were seven measurements in the memory-span tests, it is possible that one of these is more reliable than another. The rank of the individuals in each test in a given measurement was correlated with their rank in the same test in each of the other measurements. For example, the memory-span for letters was correlated with the amount of error in letters, with the average size of the largest group reproduced in letters, and with the average number of positions correctly reproduced in letters. The 'First Excess of $1/3$ Error' and 'The First Excess of $1/2$ Error' were not considered in the correlation. Only the performances of Group 1 in Test I. were considered. Tables III. and III. A show the results of these computations. In Table IV. are shown the correlations

TABLE III

SHOWING THE CORRELATION, R , BETWEEN THE DIFFERENT MEASURES OF THE PERFORMANCES IN THE MEMORY-SPAN TESTS, GROUP I, TEST II.

$$R = 1 - \frac{6\Sigma g}{n - 1}$$

Tests

Measures	Letters	Numbers	Syllables	Words	Vocabularies	Poetry	Prose	Ave.
Correlation, R , of Memory-span with								
Amount of error.....	.05	.48	.14	.48	.91	.52	.48	.44
No. of members.....	.24	.29	.53	.48	1.00	.22	.22	.43
Size of groups.....	.64	.53	.48	.48	.91	.48	.68	.60
Positions.....	.43	.48	.38	.22	.91	.31	.48	.46
Amount of error with no. of members.....	.14	.81	.34	.34	.91	.81	.34	.53
Size of groups.....	.14	.53	.15	.71	1.00	.91	.71	.59
Positions.....	.24	.72	.24	.64	1.00	.81	.67	.62
No. of members with size of groups.....	.23	.48	.34	.43	.91	.72	.53	.52
Positions.....	.34	.62	.34	.14	.91	.81	.57	.53
Size of groups with positions..	.22	.48	.81	.53	1.00	.91	.67	.66
Ave.....	.24	.54	.38	.45	.95	.65	.54	

between each test in a given measure and the same test in other measures. In Table III. A the averages of these seven correlations between any two measures are shown. The general averages given at the bottom of Table III. A may be considered as giving the relative reliability of the five measures

TABLE IIIA

SHOWING THE AVERAGE CORRELATIONS BETWEEN THE DIFFERENT MEMORY-SPAN MEASURES AS COMPUTED FROM ALL THE TESTS, THE INDIVIDUAL CORRELATIONS OF TABLE III

	Memory-span	Amount of Error	No. of Members	Size of Groups	Position
Memory-span.....		.44	.43	.60	.46
Amount of error.....	.44		.53	.59	.62
No. of members.....	.43	.53		.52	.53
Size of groups.....	.60	.59	.52		.66
Positions.....	.46	.62	.53	.66	
Ave.....	.48	.55	.50	.59	.57

that were correlated. If so; the Size-of-the-Groups measure is the most reliable, having a correlation of .59; and the Memory-span measure is the least reliable, having a correlation of .48. In the same way, the averages at the bottom of

Table III. may be considered as giving the relative reliability of the seven memory-span tests, and, if so, the test with vocabularies is the most reliable, having a correlation of .95; and the test with letters is the least reliable, having a correlation of .24.

Having determined the reliability of the different measurements, the next step was to measure the spread of improvement from the training series by the Pearson coefficient of correlation. The rank of the individuals in the training series was determined by the average learning time of all the syllable

TABLE IV

SHOWING THE CORRELATIONS, r , BETWEEN THE AVERAGE LEARNING ABILITY, A.L.A., IN THE PRACTICE SERIES WITH EACH OF THE TESTS BEFORE (I) AND AFTER (II) THE PRACTICE

$$r = \frac{\Sigma x \cdot y}{\sigma x \sigma y}$$

A. Memory-Span Tests, with Memory-Span Measure

	Letters		Numbers		Syllables		Words		Vocabs.		Poetry		Prose	
Test.....	I	II	I	II	I	II	I	II	I	II	I	II	I	II
A.L.A..	.69	.32	.65	.47	.23	.86	.52	.15	.66	.68	.44	.82	.30	.58

B. Memory-Span Tests, with Size of Group Measure

	Letters		Numbers		Syllables		Words		Vocabs.		Poetry		Prose	
Test.....	I	II	I	II	I	II	I	II	I	II	I	II	I	II
A.L.A.....	.60	.65	.59	.60	.34	.38	.20	-.04	.51	.74	.35	.98	-.60	.66

C. Learning Tests

	12 Syllables		Diagrams A		30 Vocabs.		40 Vocabs.		Poetry		Prose	
Test.....	I	II	I	II	I	II	I	II	I	II	I	II
A.L.A.....	.91	.95	.26	-.47	.95	.77	.93	.81	.84	.77	.83	.94

rows learned in that series. Their rank in the memory-span tests was determined both by the memory-span measure, since it was the principal measure of Ebert and Meumann, and the Size-of-the-Groups measure, since it had the greatest reliability. The rank of the individuals in the learning tests was determined by the learning time since it was the only measure used. Table IV. gives the results.

The facts of Table IV. *A* show a rather high positive corre-

lation between the average learning ability for nonsense syllables and the memory-span for each of the first series of the tests. These correlations are improved in the second series of tests in the case of nonsense syllables, vocabularies, poetry, and prose; and they are reduced for the other tests.

The correlations between the average learning ability and the memory-span tests, as determined by the Size-of-the-Groups measure, which are given in Table V. *B*, compare favorably with those of Table V. *A*. There are two cases, however,

TABLE V
THE RESULTS ARE IN MINUTES, SECONDS, AND PERCENTAGES
A. Ebert and Meumann

	(1)	(2)	D_1	Per Cent. I
Learning.....	3' 22"	2' 1.6"	1' 20.4"	59.7
Relearning.....	1' 24.2"	1' 6.0"	0' 18.2"	77.9

B. Test Experiment

Learning.....	8' 27.8"	5' 14.7"	3' 13.1"	63.7
Relearning.....	3' 40.1"	2' 34.9"	1' 5.7"	74.8

where the relationship between Tests I. and II. changes direction, namely, in the tests with letters and numbers.

The correlations of the average learning ability in the training series with the first series of learning tests are on the whole higher than those with the memory-span tests, which indicates a closer relationship between these functions. These correlations are increased in the second series of learning tests for nonsense syllables and for poetry. The rank in the case of the first was determined by the twelve-syllable test series alone, since it had the same quantity as those in the test series. The most significant gain is made with this test, thus confirming the result of Table II., but in case of the diagrams, a negative correlation of .47 appears in the second test as against the positive correlation of .26 in the first test. This drop of .73 means that the individuals who did the worst in the training series did the best in the second test with the diagrams. But as measured by the learning time in Table II., more gain was made in this test than in any other one. If

the negative correlation has any significance, it certainly means that the improvement as measured by the time is not due to the training series but to the discovery of some special trick by which these nonsense visual diagrams were learned with ease. Where the measurements of Tables I. and II. contradict those of Table V., it would be hazardous to express an opinion in regard to the transfer of training. Where they do not contradict each other, it may be safe to do so. In the learning tests, there is only one such case that indicates a transfer effect from the training series, namely, the test with nonsense syllables. In the memory-span tests, there are two cases, if the gains of Group 1 over Group 2 are taken in an absolute sense without reference to the A.D.'s, namely the memory-span for nonsense syllables and the memory-span for vocabularies. But the only safe conclusion appears to be that improvement in learning nonsense syllables by one method improves the ability to learn nonsense syllables by another method. In any case, the spread of improvement from a special function is not general, but it is very specialized and affects only such other special functions that are very similar to the one specially trained. The result of the Test Experiment in regard to the transfer of training is that the findings of Ebert and Meumann and the conclusions based upon them, that there is a general memorial function and that special training in one function improves the memory in general, are not confirmed. Nor is their theory confirmed that transfer from one function to another is in proportion as the functions are allied. If the correlations in Table IV. are measures of the relations between the learning ability for nonsense syllables and the special functions experimented upon in the test series, they mean that the specially trained function is closely related to practically all of the functions measured in the test series, and that the relationship of this function to the learning functions is closer than to the memory-span functions. Either the learning and memory-span functions condition each other or both sorts are conditioned by some unknown general ability. Excepting the diagrams, the specially trained function has a relationship with all of the learning functions of the learning

tests above .75, but there is evidence of transfer to only one of these and evidence of interference with two of them. The probable meaning of this is that the conditions of improvement and of transfer are peculiar to the function especially trained and do not apply to any intrinsic relationship between the abilities of the various mental functions. What these peculiar factors are is difficult to say, but the writer has some grounds for believing that they are the special associative connections developed in the course of practice upon a given material, and in proportion as these peculiar associative connections can be applied to other materials, there is a transfer of improvement.

The next question is whether the results of Ebert and Meumann and the results of the test experiment are comparable. Did the subjects of the test experiment have as much to transfer as those of Ebert and Meumann: that is, did they improve as much? If they did not, the results of the test experiment could not be considered as modifying the results of Ebert and Meumann. We may compare the amounts of improvement attained in the two experiments by the following methods: 1. We may calculate the average learning time, (1) of the first two-syllable rows learned by each method in the training series, the same (2) for the last two by each method, and then take the ratio, per cent. 1, of the last to the first as the measure of the amount of improvement. 2. We may take the absolute difference, D_1 , between these two learning times as the measure of the amount of improvement. 3. We may take the ratio, per cent. 3, of the average learning time of the last two days, (3) of the practice period to that, (4) of the first two days of the practice period. 4. We may take the absolute difference, D_2 , between these two as the measure of the amount of improvement. 5. We may take all of these together. The writer has calculated all these measures from the data of the two experiments in question. The results are given in Table V. Table V. *A* shows the group averages for Ebert and Meumann's Subjects while Table V. *B* shows them for Group I of the test experiment.

It will be noticed that the amount of improvement as measured by the ratio of the last two of each method to the

first two of each method in the practice period of Ebert and Meumann's subjects reduced their learning time 4 per cent. more than those in the Test Experiment, 59.7 as against 68.7; but as measured by the ratio of the average learning time of the last two days of practice to the first two days of practice, they improved 10.9 per cent. less than those in the test experiment, 60.4 as against 49.5. The reason for this difference is that Ebert and Meumann's subjects learned their series as regards method in a fixed order, so that each subject covered the four methods used by them every two days. This was not the case in the test experiment in which three or four days were often required to learn two series by each method. The ratios of the last two days of the practice period to that of the first two days should, therefore, be regarded as more reliable, and, according to this measure, the subjects of the test experiment improved more than did those of Ebert and Meumann in their first practice period. If the absolute difference between the learning times between the first two by

TABLE VI

A. Ebert and Meumann

	(4)	(5)	<i>D_s</i>	Per Cent. 3
Learning.....	3' 50.3"	2' 1.6"	1' 20.4"	60.4
Relearning.....	1' 19.2"	1' 6"	0' 13.2"	82.0

B. Test Experiment

Learning.....	10' 17.8"	4' 47.9"	5' 29.9"	49.5
Relearning.....	3' 55.0"	2' 32.0"	1' 23.0"	66.8

each method and the last two by each method, or if the difference between the learning times of the first two days of practice and the last two days of practice be regarded as the measure of the amount of improvement, then the subjects of the test experiment improved more than twice those of Ebert and Meumann. If the quantity of material learned is important for the question at hand, Ebert and Meumann's subjects each learned thirty-two series while each of those in the test experiment learned forty-two series in the average. And with respect to time, it appears that the subjects of the test experi-

ment spent more than twice as much time in practice as those of Ebert and Meumann. The learning times for the latter were obtained by multiplying the number of repetitions by twelve seconds, the duration of one rotation of the drum from which the series were read. It appears, however, that Ebert and Meumann stopped their drum while the learner made a trial reproduction, and then started it again from the syllable on which he first halted. Such a repetition would require more than twelve seconds, thus causing the above method of calculating the learning times to be too low. However, the conclusion from the discussion is that with regard to the amount of improvement and the condition of its transfer, the test experiment is fully comparable with the one of Ebert and Meumann. Not only did its subjects spend more time in practice and learn a greater quantity of material, but they also attained a greater amount of improvement both in learning and relearning than did the subjects of Ebert and Meumann in their first practice period. Other things being equal, the subjects of the test experiment should have shown more spread of improvement than those of Ebert and Meumann, but the former, unlike the latter, instead of showing that improvement in one function raises the efficiency of every other function to a large extent, showed that it principally raises the efficiency of the one specially trained and that of other special functions to a small extent, provided they are similar.

Just what the explanation of Ebert and Meumann's results must be in the light of this conclusion is difficult to say. But certain experiments of Swift, Schuyler, Thorndike, and Book have shown that improvement in relearning after several years of no practice is much more rapid than in the original learning. In view of the extraordinary amount of memorizing done in the German gymnasium, I believe that the improvement of Ebert and Meumann's subjects in their tests must be interpreted as a rapid re-acquisition of their former memorial skill learned in the gymnasium.

The second problem discussed by Ebert and Meumann was the economy of learning. In this connection, the series

of nonsense syllables of the practice period were learned by four methods as follows: *A*. This was the whole method. The series were repeated without pause from beginning to end until the first correct reproduction was reached. *B*. The part method, in which the subject first learned the first half of the series, then the second half, and finally the two halves together. *C*. The same as the *A* method, except that the subject paused one second at the end of the fourth syllable and one second at the end of the eighth syllable. *D*. The same as the *A* method except that the subject paused two seconds at the end of the sixth syllable. The essential difference between these two methods is the size of the units in which the respective series are learned. An *A* series is learned as a unit of twelve. A *B* series is first learned in units of six, and then the two units are learned together. A *C* series is learned together from the beginning but in units of four. And a *D* series is learned together from the beginning but in units of six. These four learning methods were also followed in the test experiment, but instead of the order of succession being varied in a fixed way, so that each subject learned an equal number by each, as with Ebert and Meumann, the order was varied irregularly so that in some cases a subject learned twice as many by one method as by another. The purpose of this was to discover whether the amount of practice would be a factor in determining the most economical method of learning. The group averages on the economy of learning, as the writer has calculated them from the data of both experiments, are given in Tables VII. and VIII.

TABLE VII
(EBERT AND MEUMANN)

Method	<i>A</i>		<i>B</i>		<i>C</i>		<i>D</i>	
	L.	Rl.	L.	Rl.	L.	Rl.	L.	Rl.
S's.								
Ave's.	10.5	4.4	9.2	4.7	7.9	4.1	9.4	4.4
L. + Rl. of ave's.		14.9		13.9		12.0		13.9
Rl. after 24 hrs.		56.9		47.5		49.7		63.4

Commenting upon these results, Ebert and Meumann say that the *C* and *D* methods are the quickest in learning and

produce a retention of average security. The *D* method is relatively quick in learning but uncertain in retention. The *A* method is the slowest in learning but has the greatest security in retention and certainty in reproduction. The *C* method is by far the best in learning. The *A* method is equally the best in retention, and, as regards the speed of learning, the *B* method lies nearer to the *C* and *D* methods than to the *A* method. In general these statements are confirmed in the test experiment. The *C* method is the quickest and the *A* method is the slowest in learning. As measured by the Saving

TABLE VIII
(TEST EXPERIMENT)
Times in Minutes and Seconds

Method	<i>A</i>				<i>B</i>				<i>C</i>				<i>D</i>			
S's.	L.		Rl.		L.		Rl.		L.		Rl.		L.		Rl.	
Ave's.....	7	15.1	3	7.5	6	18.8	3	19.	5	49.9	2	55.2	6	48.5	2	29.4
L. + Rl.....			10	22.6			9	37.8			8	45.1			9	17.1
Per cent. saved in Rl. after 24 hrs...				59.				48.3				48.3				51.7

Method, the *B* method has the poorest retention, but the *D*, and not the *A* method, has the best retention. With respect to the speed of learning, the order from the slowest to the quickest of the different methods in both experiments is the same, namely, *A*, *D*, *B*, *C*. In relearning, the order for Ebert and Meumann is *B*, *D*, *A*, *C*; and in the test experiment, it is *B*, *A*, *C*, *D*.

An important question with reference to the most economical method of learning is how to measure economy in learning. Upon this matter, Meumann says: "That method of learning is most economical which secures a particular memorial effect or attains a particular memorial purpose in the shortest time, with the least number of repetitions, and with the minimum degree of fatigue; and this method may be regarded as the most economical only with reference to this memorial effect and this memorial purpose. Of these three determinations, the learning time measures the economy of time; while the economy of energy is measured by the number

of repetitions, and in less precise form, by the amount of fatigue. No other accurate means of measurement is at our disposal.¹ Meumann puts much emphasis upon the particular memorial effect desired. One method may be most economical for immediate retention, a second for temporary retention, a third for permanent retention. The quality and quantity of the material may also determine the most economical method of learning. And the character of the associations as well as the character of the reproduction, its fidelity, rapidity, and completeness must also be considered. With all of these factors, the economy of learning becomes an extremely complex problem, and much experimentation will be needed before all of these factors can be evaluated. But it is difficult to see why the time element would not be sufficient measure of the economy of anyone of these factors, especially if economy is to be measured in terms of efficiency. For example, if one method is quickest in learning and another is the quickest in relearning, and we wish to know which is the quickest in respect to both, we simply consider that here are two pieces of work to be done, learning and relearning, and that method by which both may be done in the least time is the most economical.

In Table VII. the average number of readings required by Ebert and Meumann's subjects for both learning and relearning by each method, have been added. In Table VIII., the corresponding times appear for the subjects of the test experiment. According to this measure, the *C* method is the most economical in both experiments, a fact contrary to the conclusions in the investigation of Steffens² and Pentschew³, both of whom made out strong cases for the *A*, or whole, method.

In the test experiment an effort was made to evaluate two factors that might influence the most economical method of learning, namely, practice and the memory-span. The subjects did not learn an equal number by each method. This afforded an opportunity for learning the effect of practice upon the economy of learning. The effect in question was

¹ 'Psychology of Learning,' p. 373.

² *Zeit. f. Psychol.* 22, pp. 321, 465.

³ *Arch. f. d. ges. Psychol.* 1903, 1, p. 417.

determined by calculating the Pearson coefficient of correlation. This was a difficult matter, but space permits only to say that it resulted in $+.14$, a figure indicating some relationship but not enough to be of importance.

To determine the correlation between the memory-span and the size of the unit into which the different learning methods naturally divided the syllable rows, the average memory-span of each subject in all of the memory-span tests, *i. e.*, those before practice were calculated. Each of these averages was then arrayed with the size of the unit in which the subject learned the quickest as in Table IX.

TABLE IX

S's.	Ave. Memory- span	Most Economical Unit
Nl.....	10.57	6
Cl.....	10.30	4
Sy.....	10	12
Gr.....	9.85	12
Jn.....	7.95	12
Bv.....	7.20	6
My.....	6.93	4
Bh.....	6.76	4
Ave.....	8.67	7.5

$r = .46.$

The Pearson coefficient of correlation as calculated from these averages is $+.46$, which shows that the size of the units in which the subjects learned is a much more important factor in the economy of learning than the amount of practice, and that it is a factor to be reckoned with in connection with this problem. The importance of this factor is still more evident when we compare the memory-span of Group 1 in nonsense syllables with the size of the most economical learning unit, as in Table X.

TABLE X

Method	Ave. L. T.	Size of Unit in Each Method	Ave. Mem. Span of Group 1 in Syllables
A.....	7' 15.1"	12	4.37 Test I
B.....	6' 48.5"	6	4.54 Test II
D.....	6' 18.8"	6	
C.....	5' 49.9"	4	Ave. 4.45
Ave.....	6' 30.6"	7	

$r = .91.$

From Table X. it will be seen that the most economical unit, 4, lies very near the average memory-span of Group 1, 4.45. The units of the *B* and *D* methods, which are 6, lie 1.15 farther away from the group memory-span. The unit of the *A* method, which is the slowest of all, lies 7.15 away from the group memory-span. The relationship between the quickest learning method of Group 1 and its most economical learning unit is beautifully expressed by the Pearson coefficient, which is $+.91$. The correlation between the optimum method and the optimum unit of the group is thus $.91$, while that corresponding to the individuals is $.46$. It is probable that the latter would approach the former if a greater number of individuals were compared, and, if so, we must consider the size of the unit in which an individual learns as the principal determinant of his optimum method. The size of this unit is conditioned by the learner's memory-span. Under these conditions, we could not advise the whole method as such, or the part method as such, or the modified whole method as such, as the optimum method, but rather that an individual should learn in such units that lie near his memory-span for the material in question.

It is possible, however, that the character of the associations may have something to do with the optimum method. It will be noticed from Table X. that the learning time for the *B* method is shorter than for the *D* method, but its relearning time is much greater, so that the per cent. saved in the *D* method is 63.4 as against 47.5 in the *B* method. Considering both the learning and the relearning times, the *D* method is more economical than the *B* method. In the *B* method, the halves of the series were first learned separately and then together, while in the *D* method the halves were learned together from the beginning. My. and Bh. reported that learning the halves of the *B* method together after having learned them separately required a new organization of the entire series. They could not put the two halves together mechanically, but had to break up the associations formed with the halves and build new ones, thus causing interference between the two sets of associations. Such a condition would

easily explain the small saving in the *B* method, for the interference would naturally weaken both sets of associations. On the same grounds, the great saving in the *D* method may be explained as being due to the absence of such conflicting associations. But the difference in the saving between the *A* and *C* methods is better explained as due to the difference in imprinting. So far as this experiment goes, the advantage of the *C* and *D* methods is due to the fact that they do not form conflicting associations but yet adjust the material to the learner's memory-span. This would suggest that a learner should generally follow a modified whole method with pauses in between subdivisions the size of which is determined by the memory-span for the material in question.

The third problem which comes up in this experiment but which was not considered by Ebert and Meumann is the influence of quantity of learning material upon the memory-span. The memory-span tests varied so widely in quantity that they afforded an excellent opportunity for studying the effect of the quantity of the stimulus upon the quality of the immediate reproduction. A study of this sort is not only interesting for its own sake, but also for the light it throws upon the problem of the economy of learning.

The quality of the reproduction was measured in four ways: (1) the average amount of error; (2) the average number of members reproduced; (3) the average size of the largest groups reproduced; and (4) the character of the position reproduced. The nature of the first three measures was explained in the first part of this report, but a little more should be said about the positions. These were divided into three classes: (1) those that were correctly reproduced with respect to their relative positions to either end of the series; (2) those that were misplaced; and (3) those that were omitted.

The measurements of the subjects' performances in a given quantity in a given test were all placed in vertical columns, and then the averages of the measurements added vertically for both Tests I. and II. irrespective of the group divisions were calculated in order to show the effect of the quantity of the stimulus upon the amount of error. The A.D.'s of the

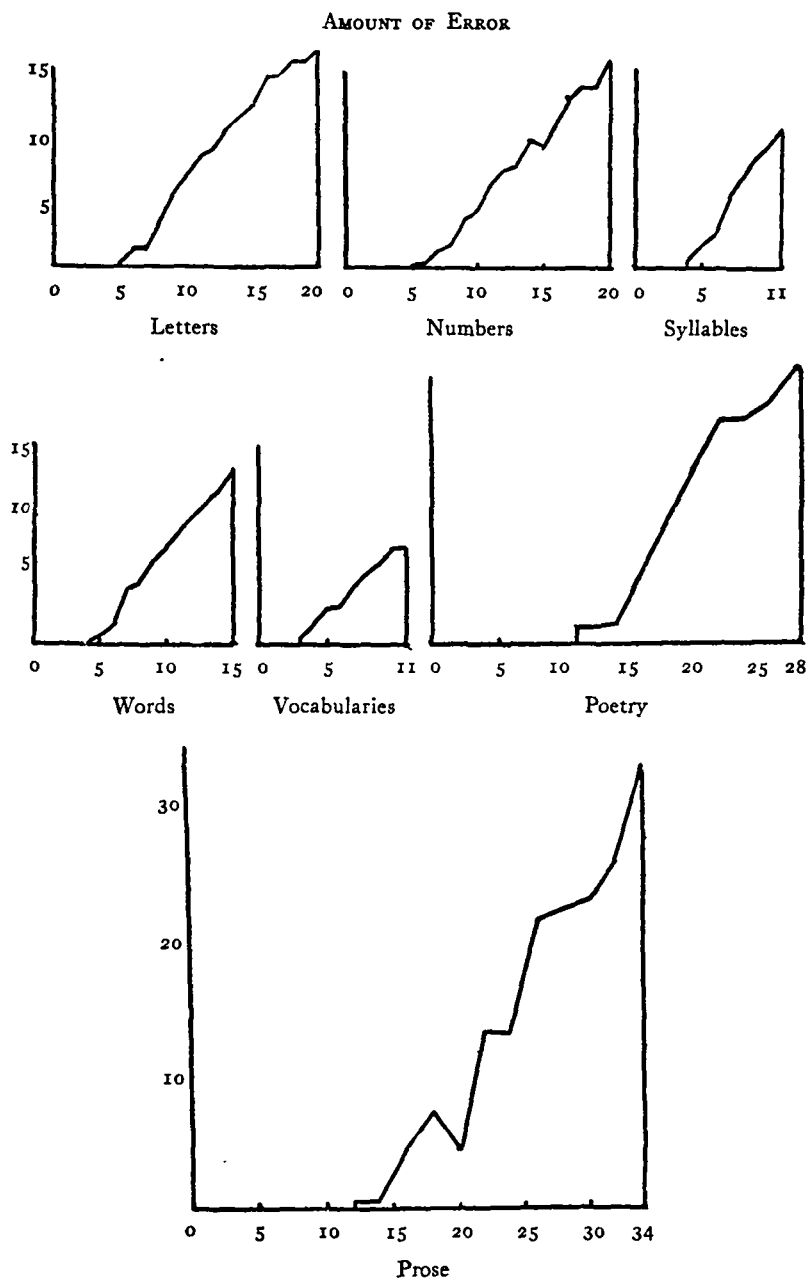


FIG. 3. Curves showing effect of quantity upon amount of error.

Ordinates = Ave. no. errors.

Abscissas = Ave. no. members in series.

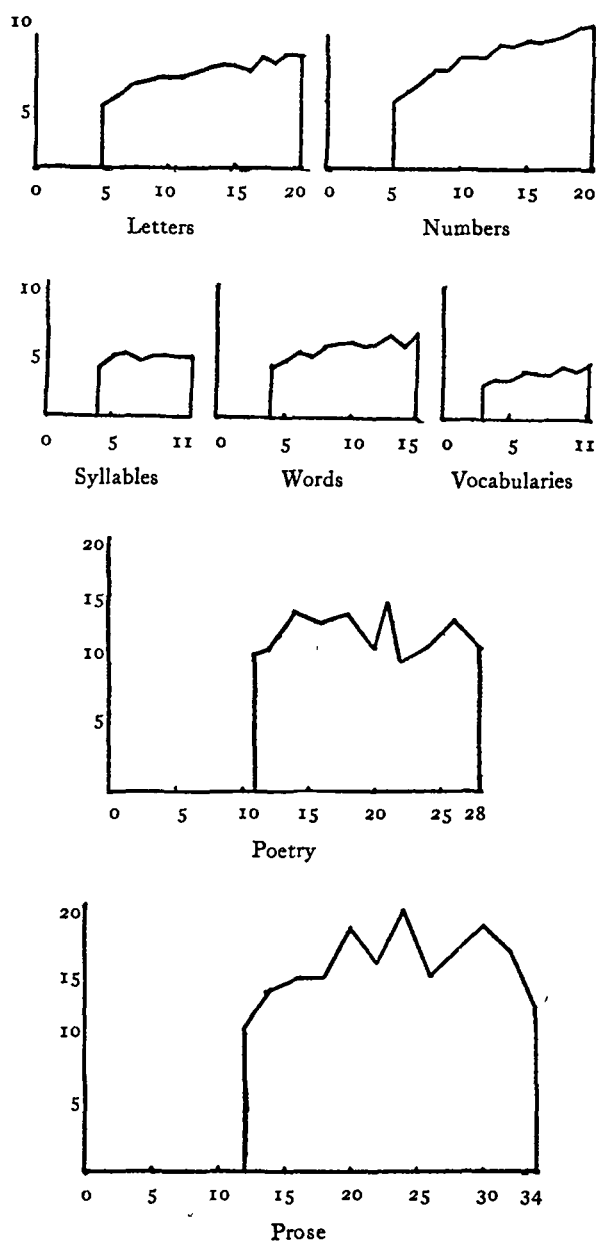


FIG. 4. Curves showing effect of quantity on number of members reproduced.
 Ordinates = Ave. no. of members reproduced.
 Abscissas = Ave. no. of members in series.

averages were also calculated. The averages in the tests with letters, numbers, and syllables are each based upon sixty measurements; and in those with vocabularies, words, poetry, and prose they are each based upon thirty measurements. The results with reference to the amount of error are presented graphically in Fig. 3. It will be noticed that the amount of error increases directly as the quantity of stimulus.

The results with reference to the number of members are presented graphically in Fig. 4. Perhaps the effect of quantity upon the number of members could be expressed by saying that increasing the stimulus beyond a quantum a little larger than the memory-span increases the number of members reproduced by a fraction whose value approaches zero. In fact, it is not very incorrect to say that beyond such a point the number of members reproduced is not increased. For example, in the case of syllables increasing the number of members beyond six does not increase the size of the reproduction beyond the average memory-span, 4.45. In case of prose, the number of members is not increased materially for sentences having more than 20 words, the average memory-span being a fraction over seventeen.

Fig. 5. shows the effect of quantity upon the size of the groups correctly reproduced. It will be noticed that increasing the quantity of the stimulus beyond the memory-span decreases the size of the groups correctly reproduced. This is more noticeably true of meaningful material than of nonsense material. For example, when a sentence of twenty words is read, the average size of the largest groups reproduced was 14.43 words. When a sentence of thirty-four words was read the average size of the largest group was only 4.75. When we consider that the ability to comprehend long spoken or written sentences depends not upon the ability to get a word here and there, but upon our ability to get a connection of all the words, we see the pedagogical importance of the above relationship between quantity and an immediate reproduction that is held responsible for order. This is all the more significant because most things that are read and heard are given only one repetition.

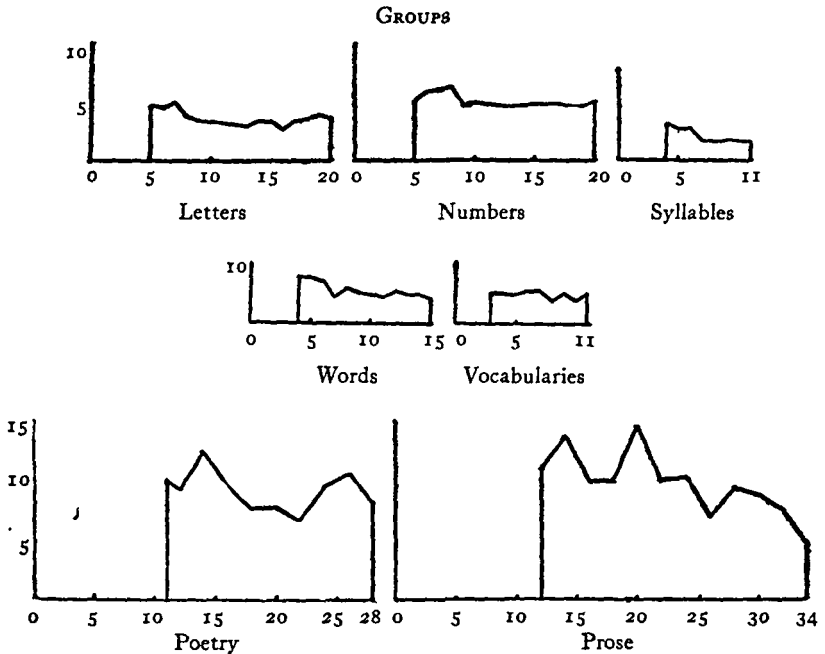


FIG. 5. Curves showing effect of quantity on size of groups.
 Ordinates = Ave. size of groups reproduced.
 Abscissas = No. of members in series.

The effect of quantity upon the character of the positions reproduced and omitted is shown in Fig. 6 which depicts those from the test with disconnected words. The curves here are typical of all the tests. They may be described verbally as follows: If in coördinates, the frequency of positions reproduced is plotted along the ordinates and the successive positions are placed along the abscissa line, the distributions of the positions are as follows: First, for those correctly reproduced: When the stimulus is less than the memory-span, their distribution is a rectangle. As the stimulus increases beyond this point, the roof of the rectangle begins to fall, first in the middle and then along the entire line, and, as it does this, the sides of the rectangle become a little shorter, but yet remain quite high even for large quantities, the side standing for the end of the series being higher than that for the first of the series. The roof of the rectangle thus tends to form a curve

which has the reverse shape of the probability curve. Second, the positions omitted: The curve for this omission is just the reverse of the one described. The distribution for quantities

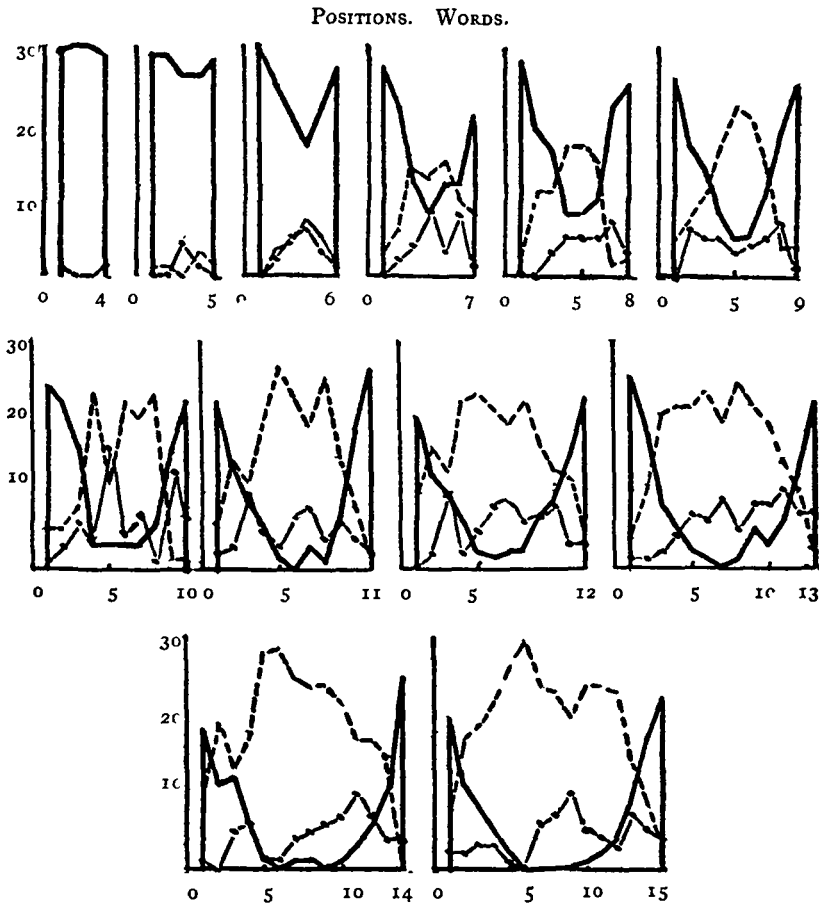


FIG. 6. Curves showing effect of quantity on distribution of positions correctly reproduced.

Continuous Lines = positions correctly placed.

Dots and dashes = positions misplaced.

Dashes only = positions omitted.

Ordinates = frequencies of reproductions.

Abscissas = order of positions in each series.

less than the memory-span is a zero abscissa line, and, as the quantity increases, this line forms a little lump in the middle

which grows larger and larger as the stimulus is increased, until the high point embraces all the positions in its ordinate; and from this, the line gradually falls toward the base at each end. Third, the misplaced positions: The distribution of these is irregular, but their tendency is to follow omissions.

The essential significance of these conclusions on the effect of quantity upon immediate reproduction is that the mind in learning a quantity of material must proceed by steps just as the body does in covering a quantity of space. The extent of that step is fixed just as much as the bodily step, and just as the body falls flat when it tries to take too large a step, so the mind literally falls when it tries, in one act, to comprehend a material that is much beyond its memory-span. It simply does not get it, and, in fact, it gets less than if it proceeded by small steps.

The significance of this for the economy of learning is evident. It means that the mind can learn only by parts, no matter what the method of reading is. If the whole method is followed, it means that different parts come above the memorial threshold at different times. If the part method is followed, fewer readings will be required to bring the part above this threshold. However, a correct reproduction must reproduce not only the members but also the order of the series. We have seen that learning in quantities greater than the memory-span reduces the learner's capacity for this kind of reproduction. Is it not then a saving of energy and time to adjust the optimum learning unit to the learner's memory-span? It appears that the above facts are a fatal blow to the much puffed-up whole method as such, and that the part method is much more in agreement with psychological laws. If so, the economy of learning as regards method is how to avoid conflicting associations.