Vol. XXXI No. 3 *PSYCHOLOGICAL REVIEW PUBLICATIONS* Whole No. 142 1922

Psychological Monographs

EDITED BY

JAMES ROWLAND ANGELL, Yale University.
HOWARD C. WARREN, Princeton University (*Review*)
JOHN B. WATSON, New York (*J. of Exp. Psych.*)
SHEPHERD I. FRANZ, Govt. Hosp. for Insane (*Bulletin*) and
MADISON BENTLEY, University of Illinois (*Index*)

STUDIES FROM THE PSYCHOLOGICAL LABORATORY OF THE UNIVERSITY OF CHICAGO

The Conditions of Retention

BY

C. W. LUH

Professor of Psychology, Southeastern University, Nanking, China

PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.

Agents: G. E. STECHERT & CO., London (2 Star Yard, Carey St., W. C.)
Paris (16 rue de Condé)

ACKNOWLEDGMENT

To Prof. H. A. Carr I wish to express my deepest sense of obligation for constant supervision and encouragement. Thanks are also due to my friends and fellow-students whose patient service made this study possible. Dean J. R. Angell and Dr. J. R. Kantor took great interest in my work. To them I owe much that cannot be conventionally acknowledged.

CONTENTS

I. INTRODUCTION

The problem of the present study is to investigate the nature of the curve of retention under certain variable conditions. The conditions that can be independently varied are numerous. Only two series of experiments have been systematically carried out:

(1) Varying the methods of measuring the amount of retention and

(2) Varying the degree of mastery in the original learning.

As established by Ebbinghaus,[1] the curve of retention for nonsense syllables drops very rapidly during the first 20 minutes after learning. More than half of the original material is lost at the end of the first hour. The subsequent fall of the curve becomes less and less abrupt until about 1 day, when the curve runs almost parallel to the abscissa. Similar but far less extensive experiments were performed upon himself with meaningful material, *i. e.*, poetry. From these results Ebbinghaus deduced the following equation of the curve which has remained a classic of forty years' standing:

$$b = \frac{100\,k}{(\log t)^c + k}, \text{ or } \frac{b}{100 - b} = \frac{k}{(\log t)^c}$$

In this formula b = percentage of retention, $100 - b$ = percentage of forgetting, t = length of the interval in no. of min., and k and c are constants.

With Müller and Schumann,[2] the technique was greatly improved in the construction and the presentation of the syllable series. Their results corroborated those of Ebbinghaus, though the amount of forgetting was less than previously reported.

[1] "Ueber das Gedächtniss: Untersuchungen zur experimentellen Psychologie," Leipzig, 1885. Eng. Tr. by Ruger and Bussenius, Teachers' College, Columbia University, 1913.

[2] "Experimentelle Beiträge zur Untersuchung des Gedächtnisses," *Zeitschr. f. Psychol.*, Vol. VI, 1893.

It was in 1903-04 that Radossawljevitch[3] attempted a more careful determination of the curve on a much more extensive scale. He employed altogether 29 subjects, as against Ebbinghaus' one, who was himself. Availing himself of the improved technique of Müller and Schumann, he further introduced accent and rhythm into the act of learning. On the whole, the curve he described conforms to the Ebbinghaus type. The amounts of forgetting after relatively short intervals were less than either Ebbinghaus or Müller and Schumann discovered.

Finkenbinder,[4] who re-attacked the problems in 1912, first used the *anticipation* method in presentation. By that method the subject was required to anticipate each syllable by pronouncing it aloud within the exposure period of the syllable preceding. Successful anticipation of the whole series constitutes the standard of learning. Altogether eleven different intervals were used, and these were carefully distributed so as to eliminate diurnal variations. He summarized his data as follows: "The curve of forgetting for nonsense syllables in series of twelve, as determined by the lapse of time, is a uniformly progressive curve much as Ebbinghaus found, but under the conditions of our investigation, the progress of forgetting is slower than Ebbinghaus found it to be, but somewhat faster than Radossawljevitch found."[5] Then from his numerical data he found this astonishingly simple equation: Forgetting $=$ 10 (log. of time in No. of min. $+$ 1), and apologized for its simplicity.

In the above quoted investigations, the method for measuring the amount of forgetting was the relearning or "saving" method. Forgetting is expressed as a quotient of the amount of time for relearning and for the original learning. Strong[6] first undertook

[3] "Das Behalten und Vergessen bei Kindern und Erwachsenen," Leipzig, 1907.

[4] "The Curve of Forgetting," *Amer. Jour. Psychol.*, Vol. XXIV, 1913.

[5] *Ibid*, p. 32. Finkenbinder used series of 12 syllables each; Ebbinghaus, 13; Radossawljevitch, 8, 12 and 16. Finkenbinder should have compared his results with that part of R's data which were derived from 12-syllable series only.

[6] "The Effect of Time Interval upon Recognition Memory," *Psych. Rev.*, Vol. XX, 1913, pp. 334 ff .

to extend the investigation with other methods of measuring results. The method used in his experiments was the method of recognition, or selection, and for material he constructed series of common English words of 40 each. Out of each 40, a second series of 20 words was drawn at random, which latter served as the exposure list. The first list of 40 was given the subject for recognition after the lapse of a designated interval of time. From these the subject was required to select the exposure list. The experiments included 14 intervals and 5 subjects, but only 15 measures for each interval. The curve of retention thus determined was similar in shape to that of Ebbinghaus. The two almost agreed as to actual amounts of forgetting until after the lapse of the 4-day interval. Strong, therefore, concluded that "there is no difference in the form of the curve for retention in recall and recognition memory." By this he did not specify whether he implied the logarithmic formula, or only the general shape of the curve as to its initial drop and negative acceleration.

Bean[7] thought that the Ebbinghaus tradition could be substantiated on more extensive grounds. In his experiments two distinct kinds of material were used, (1) series of 9 consonants of the English alphabet, and (2) typewriting. In (1) he used two methods for the measurement of retention. A. The method of written reproduction, in which the subjects were allowed 2 min. to reproduce the series of consonants which had been learned from 1 to 28 days previous to recall. B. The method of selection and reconstruction, in which the subjects spent 90 secs. in selecting the original series of 9 out of the total group of 18 which was being presented, and then took 30 secs. to put the original 9 in the right order. In (2) typewriting, Bean adopted a method similar to that used by Book in his pioneering studies in typewriting.[8] Bean's results may be summarized as follows:

The curve of forgetting may be studied (1) as to its general

[7] "The Curve of Forgetting," *Archives of Psychol.*, Vol. XX, No. 3.

[8] "The Psychology of Skill," *Univ. of Montana Publications in Psychology*, Vol. I, 1908, Bull. 53. Psy. Series No. 1.

form; and (2) as to the absolute rate of forgetting. According to him, the curves he presented differ from those of earlier investigators and among themselves, not in general form, but only in absolute initial amount. Supposing the logarithmic type of curve to be general, then an initial high amount of forgetting will lead to apparently abrupt transitions. As to the cause of the variations in the absolute or initial amount, numerous conditions may be mentioned. 1. Degree of learning. 2. Distribution or concentration in the process of learning. 3. Different kinds of material. 4. Different methods by which retention is measured. 5. Individual differences. On the whole, he would conclude that all curves of forgetting are logarithmic curves.

As we shall later develop, the methods he used for measuring the amount of retention were mostly crude and inaccurate. He did not give any measurement for the statistical validity of his data. Further, he himself did not attempt to find equations, or rather one general equation, for his own data, and we have to take his conclusions only as pious opinions. Finally, in his experiments, none of the conditions he mentioned was controlled and independently varied.

An even more sweeping generalization than Bean's was offered by Piéron.[9] He developed a formula which was believed to satisfy all the conditions of immediate and permanent memory, retention and forgetting, muscular contraction and the phylogenetic development of retentive phenomena, etc. That formula is $m = \frac{k}{t^a (\log t)^c}$, where m = percentage of "saving," t = length of the interval, and k, a and c are constants. Then putting a = 0, it takes in Ebbinghaus' curve for human retention, $m = \frac{k}{(\log t)^c}$, as a congenial member of the "family."[10] But these deductions were based on no surer foundation than a few observations upon some pond snails.

[9] "L'Évolution de la Mémoire," Paris, 1910, pp. 256-60.

[10] Notice how ingenuity is accentuated by forgetting k in Ebbinghaus' formula. This is permissible when t is very large. But E's formula is significant only when it is small.

All the above writers seem to concur with the Ebbinghaus tradition and to agree on these points.

(1) Retention decreases with time.

(2) Forgetting is more rapid at first.

(3) The curve of retention is generally uniform so that one can state it in terms of a mathematical formula. Radossawljevitch indeed located a sudden deviation at the end of the 8 hr. interval, but that was smoothed out in Finkenbinder's experiments after the elimination of diurnal variations.

(4) There is a question as to whether

A. The curves approximate each other so closely that one can regard them as being chance deviations from some ideal curve, or

B. They belong to a single family with one or more variable parameters which assume different values according to different conditions. Strong inclines toward the former alternative, but Bean explicitly favors the latter.

It is to be noticed that conclusions 1 to 3 are but descriptive summary statements of the facts as observed under their particular experimental conditions, while 4 is a mathematical deduction. By "the Ebbinghaus type of curve," one may simply mean that facts can be so described and graphically represented, without implying that they satisfy any kind of an equation. Neglecting all mathematical complications, may not the phenomena of retention and forgetting be, *under all conditions,* much as Ebbinghaus described them?

In the light of these conclusions, Ballard's[11] results become significant. He experimented on school children with Latin nouns, nonsense poetry, geometric diagrams, nonsense syllables, prose, material for logical memory, but above all with ballad poetry. For the vast majority of his subjects, the material was not completely learned. Written reproduction was required after the lapse of a certain interval. He found that the highest proficiency of retention, as measured by the amount of repro-

[11]"Obliviscence and Reminiscence," *Brit. Jour. Psych. Monograph Supplements,* Vol. I, 1913, No. 2.

duced material, was reached not immediately after learning, but, as a rule, at the end of two days. The graphs in Fig. I are illustrative.[12] Both A and B are curves of retention for children about 12 years of age. The amount remembered immediately after learning is taken as the basis on which the percentage repro-

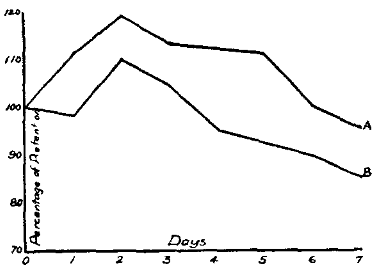


Fig I

duced later is calculated. Graph A. was based upon "The Wreck of the Hesperus," with 20% of the material remembered in the "primary test." B. was based upon "The Ancient Mariner," with 40% of the material remembered.

These results of Ballard directly contradict every one of the conclusions reached by the previous writers. Yet his data seem to be at least as valid and reliable as those of our earlier investigators in forgetting.

This apparent dilemma suggested to us what was, in a way, anticipated by Bean. When the conditions of learning and recalling are changed, not only will the "absolute" amount of forgetting change, not only will the mathematical formula change, but the phenomenon of negative acceleration may also disappear. The conditions under which Ballard experimented were greatly different from those of other investigators. The following were the most important:

(1) The age of the subjects. Ballard's subjects who manifested these peculiarities were all children. Compared with his own work upon adults, the results indicate that the curve is a

[12] *Ibid*, p. 5.

function of the age of the subjects. Other investigators worked mostly with adults.

(2) The kind of material. Ballard used but little nonsense material and did not work on a separate curve of retention for nonsense syllables. The Ebbinghaus type of curve is a curve for nonsense syllables *par excellence*. Ebbinghaus' and Radossawljevitch's experiments with meaningful material were only supplementary to their principle problem and were haphazardly performed.

(3) The degree of learning. In Ballard's experiments, the material was only partly learned. The degree of learning for the curves quoted above was 20% and 40%. The subjects had to stop at a given time limit. In all previous work, except that of Strong, the material was learned at least to the first errorless recitation.

(4) The method of measuring the amount of retention. Before Ballard, the method that was generally used was the "saving" method. Bean and Strong tried the "selection" method, but not very extensively. Ballard, however, found the written reproduction method to be more suitable for group experiments.

This comparison suggests the question as to *whether the curve of forgetting, or of retention, is a function of these conditions, or more specifically, whether each condition determines a special curve.* The present study is an attempt to answer this question. This can be carried out only by varying the above mentioned conditions independently and systematically. Unfortunately, Ballard's technique has necessitated a program which demands more time than we can at present afford. We have, therefore, limited the scope of the problem by taking up only the third and fourth factors, leaving the first two to a later investigation.

General Description of the Experiments

The experiments were performed in the Psychological Laboratory of the University of Chicago, the first series from May to August, 1919, and the second from October, 1919, to February, 1920.

(1) *Material*

Series of nonsense syllables of 12 each were used. With the English alphabet, a list was made of all possible combinations of two consonants joined by a vowel in the middle, except those which end in y. From this list we eliminated all the English, French, German and Latin words. The revised collection was then submitted separately to four graduate students of the department who checked every syllable which happened to call up immediate meaning associations. All syllables thus marked out by more than one of the observers were further eliminated from the final list.

In the construction of the series, Müller's[13] rules were observed as closely as possible. That is,

(1) All the initial and final consonants of the same series are different.

(2) Since we did not resort to the use of diphthongs, we had five vowels as against Müller's twelve. The terms were arranged so that no two of four consecutive syllables have the same vowel.

(3) No two consecutive syllables have any consonant in common.

(4) No group of consecutive syllables constitutes a polysyllabic word or a phrase.

Thus we have improved upon Müller in at least two respects. (1) He overlooked monosyllabic words, which, perhaps, is not so serious an omission in German as in English. (2) While he allowed a syllable to end with the initial consonant of the preceding syllable, we excluded all such cases.

Since in our investigation we required the subject to spell each syllable letter by letter, instead of pronouncing it as a whole, we are no longer concerned with Gamble's rules[14] which were formulated as safeguards against the inherent defects of English orthography. We also think that spelling the syllable letter by letter tends to minimize its meaning associations.

[13] *Op. cit.*, p. 106.

[14] "A Study in Memorising by the Reconstruction Method," *Psych. Rev., Monograph Supplements,* Vol. X, No. 4, p. 2.

Altogether about 90 series were constructed so that a subject could serve extensively without resorting to learning a single syllable twice.

(2) *Apparatus*

The apparatus was an ordinary rotating drum used in the Chicago laboratory for most of the memory experiments. After the series were typewritten on strips of white manila card, they could be easily fixed to the drum. One syllable was exposed at a time through an aperture in the screen attached to the posts of the drum. It was found more convenient to run the apparatus by hand than by a mechanical device, since the experimenter had to keep his eyes on the aperture in order to be perfectly sure that the subject spelled the syllable completely before the succeeding one was exposed.

In experiments like these, every moderately loud noise may be disturbing. For this reason the experimenter had to keep the time by running a telephone wire from an adjacent room in which a metronome was set at two seconds.

In the later part of the first series of experiments and throughout the second, it was thought worth while for purposes other than that of the present study to keep minute records of every correct or incorrect response. A Remington typewriter No. 6, invisible, served as the apparatus, so that the subject would not be distracted by what was being recorded. Two keys were arbitrarily chosen to mark success and failure. The striking of the key made a noise every two seconds. The effect of this apparently disturbing factor was inappreciable, so far as we can determine from the practice curves of the individuals.

(3) *Method of Presentation*

The subject was seated at a convenient distance in front of the rotating drum. Before the presentation of every new series, the experimenter gave the signal "ready," one second after which the first syllable was exposed. This signal became superfluous after one or two weeks of practice. The time of exposure for each syllable was 2 secs. No restriction was made as to the method of learning, excepting that the subject was warned not to form artificial meaning associations. Beginning with the second exposure of the series, he was instructed to attempt to anticipate each syllable by spelling it aloud before it was exposed. Usually 6 seconds were allowed between successive presentations of a series, during which the experimenter shifted the drum and made the necessary records. A series was considered learned once the subject successfully anticipated every syllable. All series were learned by successive presentations in a single sitting.

Each subject was required to return at the same time of the day, but could skip two or three days in succession between series. Only one series was learned in a day. Occasionally the learning of a new series followed immediately after the recall of the preceding one. But this never happened in the second series of experiments.

The different intervals and the methods of measurement, as will be described later, were distributed according to a tentative scheme drawn for each subject. Sometimes that schedule had to be slightly changed, but care was always taken to reduce to a minimum the effect of uncontrolled practice.

Each subject was given 4-6 series for preliminary practice. These results were not counted. (Subject C, who learned only 20 series, began to work late in the summer. Only 2 practice series were possible with this subject His results could have been improved by giving 2 or more additional preliminary series).

(4) *Subjects*

Ten subjects a day were all we could handle. We had two groups of ten each. The first group included one instructor and three graduate students of the department and six Chinese students of the University who had one time or another taken some work in psychology. In the second group, there were six graduate students and one senior of the department, including the experimenter. The other three were Chinese students of the University, two of whom never had any work in psychology. Subject Y, a Chinese student who learned 20 series in the summer, served again as one of the ten. When the experimenter served as subject, the series were given by Mr. T. L. Wang, who had also served in previous experiments. In this case, the series were made according to rules unknown to the experimenter. They were very similar to the ordinary ones.

Apparently the Chinese students had no serious difficulty in learning this type of material. They were able to memorize directly without translating the exposed material into Chinese equivalents. On the whole, they learned the series very much faster than did the Americans.

II. THE AMOUNT OF RETENTION AS A FUNCTION OF THE METHOD OF MEASUREMENT

In the first series of experiments, two methods of recall were used, which together furnished five measurements of the amount of retention.

1. The Anticipation and Relearning Methods. In one half of the series given to each subject, the method of presentation in relearning was identical with that of the original learning.

(A) Thus, the subject was required to anticipate the series at the very first presentation, at the rate of 2 sec. for each syllable. The number of correct syllables was recorded. That number, expressed as a percentage of the whole series, established a measurement of retention in terms of anticipatory recall.

(B) After the first record was taken, the series was exposed as many times as necessary for complete relearning. A measurement of retention was thus furnished by the "Saving Method."

2. The Reproduction, Recognition and Reconstruction Methods. These methods were used with the other half of the series given to each subject.

(C) The subject was first furnished with a recall blank, on which were three columns of figures from 1 to 12, each with a blank space to the right. He was instructed to write down the original series in the right spatial order, beginning with the left hand column, but not necessarily in the same temporal order as the series was learned. At the end of 1 min. the experimenter gave a signal, at which the subject began to write in the middle column, filling out spaces that had been left open during the first minute, or correcting any mistake that he thought had been made. Another signal was given at the end of the second minute. At this, the subject changed to the right hand column. Three more minutes were allowed for further reproduction and correction. The subject could, of course, "give up" at any time, or finish the whole series before the lapse of the first or second minute. A time record was also taken for each series completed within the above 5 min. This constituted a measurement of retention by the Written Reproduction method.

(D) Immediately following upon written reproduction, a group of 24 syllables was given the subject, out of which he was required to select the original 12, no more and no less. At the end of 90 sec. the experimenter quietly took a record of the number of correct and incorrect syllables selected up to that time. A similar record was taken when the subject completed the selection of 12 syllables, correctly or incorrectly, together with a time record. The time limit, 90 sec., was determined from the averages of the preliminary records of all the subjects in the first series of experiments. This process gave a measurement of retention in terms of Recognition.

(E) Finally, the subject was furnished with the original 12 syllables on separate slips of white manila card and was required to reconstruct the order of the series. The actual order of reconstruction was recorded, and also the amount of time spent in the reconstruction process. This we may call the Reconstruction or Rearrangement method of recall.

Intervals of Time

Five intervals of time, *i. e.*, 20 min., 1 hr., 4 hrs., 1 day, and 2 days, were selected with two considerations in view.

(1) To facilitate comparison with earlier reports, it was necessary to fix our program into that of other investigators. Accordingly, intervals which had not been included in the work of one or more of our predecessors do not appear in our plan. We further took into consideration whether the points on the curve to be thus empirically determined would be likely to represent an equation, if there be one. Cf. Table VII.

(2) We also tried to avoid the effect of diurnal variations. Later we found that the 4 hr. interval was too long for our purpose. Any interval longer than 4 hrs. and shorter than 24 would be too much beyond our control.

Methods of Scoring

(1) Anticipation. With this method complete retention is the successful anticipation of every syllable. There are 12 syllables in each series. On the basis of 12 one can easily convert an actual score into percentage terms.

(2) Relearning. The usual "saving method" was used, in which the number of presentations in relearning minus one, divided by the number of presentations in learning minus one, multiplied by 100, gives the percentage of forgetting.

(3) Written Reproduction. First, the reproduced amount is compared with the possible amount. X correct syllables reproduced is scored at X/12, which, multiplied by 100, gives a percentage score. A syllable with only 2 letters correct is scored ½, as is also one with the initial and final consonants inverted.

Secondly, we took into consideration the position and sequence of each reproduced syllable. The difficulties for such minute scoring are two. (1) The relative value of the reproduced amount on the one hand and of position and sequence on the other can only be arbitrarily determined. (2) By chance the subject may reproduce a certain number of syllables in the original position or sequence without actually remembering either, and it is well nigh impossible to score this chance factor. In the present study, the total amount of reproduced material will be roughly scored ½, and position and sequence ¼ each. It is assumed that material contributes as much value as position and sequence put together, and the latter are again considered to be equal in value. The chance factor is neglected. The records are, therefore, scored too high, but the extent of this effect one can easily approximate when we come to deal with the method of scoring reconstruction. Any such arbitrary process will, of course incur all the criticism that has been heaped upon Lyon[1] by writers like Kjerstad.[2] The latter, however, also neglected the chance factor. Our interpretations will be based mainly upon

[1] "A Rapid and Accurate Method for Scoring Nonsense Syllables and Words," *Amer. Jour. Psychol.*, Vol. XXIV, 1913, pp. 525-31.

[2] "The Form of the Learning Curves for Memory," *Psych. Rev., Monograph Supplements,* Vol. XXVI, 1919, No. 5, pp. 14 ff.

the first method of scoring, which does not take into account position and sequence.

Further, unlike the other methods of scoring, the values for written reproduction are independent of the amount of error made in recalling. The amount of error may also be reduced to percentage scores when compared with the actual or the possible amount of retention, *i. e.*, the amount of error may be computed as a percentage either of the whole series or of the reproduced material only.

(4) Recognition. When a number of syllables originally learned are mixed up with an equal number of new ones and then presented to the subject for recognition, the outstanding fact is that, by pure chance, one will most probably draw half of the original ones. Bean[3] overlooked this difficulty. Consequently, all his values were above 50% even to the end of the 28th day.

To eliminate this chance factor, Strong[4] devised the formula,

$$\text{Retention} = \frac{\text{Correct recognitions}}{\text{Total no. presented}} \times \frac{\text{Correct} - \text{incorrect recognitions}}{\text{Correct} + \text{incorrect recognitions}} \times 100.$$

No doubt, this formula takes into account the extremes of probability. That is, out of X things learned, a recognition of X/2 is scored 0, while a recognition of X is scored 100. Beyond that, the formula is exposed to numerous difficulties and seems to defeat its own purpose. The author pointed out that "it penalizes mistakes a little more than is warranted on a basis of chance." As a matter of fact, it penalizes sometimes too much and sometimes too little. In other words, the scores given on this basis are not always proportional to the probability. That proportionality, it seems to me, ought to be the criterion for the validity of the formula. For instance, $\frac{8}{24} \times \frac{8-1}{8+1} \times 100 = 25.9\%$. Similarly, 8 correct ones and 2 incorrect ones would be scored 20.0%, and 8 correct ones and 3 incorrect ones 15.6%. Now

[3] *Op. cit.* His reproduction method is not any better. When one is required to write down 9 out of 18 consonants it is most probable that 4 or 5 will be correct, even though one does not realize there are *only* 18 consonants in the English alphabet.

[4] *Op. cit.* p. 355. Cf. *Psych. Rev.*, Vol. XIX, pp. 457 ff.

when 24 members are presented, the respective probabilities for these three combinations to occur are $\frac{945}{52003} : \frac{3465}{52003} : \frac{9075}{52003}$, provided that, by pure chance, one is as likely to take 9 as 10 or 11. It is very difficult to see how the chance factor is counter-balanced by giving scores such as the above.

Again, suppose that of the total group presented, the subject selects only 1 and that 1 be correct. According to the formula, this performance would be scored $\frac{1}{\text{Total no.}} \times \frac{1-0}{1+0} \times 100.$

But if in the total number presented, the number of original ones is equal to the number of new ones, the subject will be just as likely to draw a correct one as an incorrect one. Such a performance should be scored 0.

Other defects of the formula, while not inherent and unavoidable, result from assumptions which one makes in applying it. Thus Strong required his subjects to classify their judgments according to degrees of certainty. After the first class, *i. e.*, the most certain one, is scored by his formula, the second and third classes cannot be penalized as rigidly, since chance has been greatly reduced by the exclusion of the correct ones in the first class.

Further, on the basis of the first class as 1, Strong scored the second class ¾, and the third class ¼. These values are entirely arbitrary and have nothing to do with the formula. But this method of scoring, together with the last named oversight, certainly helped to make his curve of recognition memory something like Ebbinghaus' curve.

In the formulation of our own method of scoring, we first take it for granted that, if there were no chance factor, each score should then increase upon the next by a constant amount.

Then we calculated the probability of each kind of combination. Thus, when the total group of 24 is presented, there are 2704156 possible ways to take 12. These are classified as follows:

No. in which there are 12 correct and	0 incorrect,	1
11	1	144
10	2	4356
9	3	48400
8	4	245025
7	5	627264
6	6	853776
	etc.	

Six correct is the highest probability, the combination which is most likely to occur on the basis of pure chance. If we next regard this probability as 100% chance, we have

6 correct and	6 incorrect,	100.00% chance,
7	5	73.47%
8	4	28.70%
9	3	5.67%
10	2	.51%
11	1	.02%
12	0	.00%

Everything below 6 correct may be disregarded. Now divide 100 into 6 equal intervals, from 6 to 12, for scores when the chance factor is not deducted. Deducting from each interval the relative amount of chance, the final scale is

TABLE I

CORRECT	INCORRECT	PRELIMINARY SCORE	CHANCE	FINAL SCORE
6	6	0.00	100.00	0.00
7	5	16.67	73.47	4.42
8	4	33.33	28.70	23.76
9	3	50.00	5.67	47.17
10	2	66.67	.51	66.33
11	1	83.33	.02	83.31
12	0	100.00	.00	100.00

By extending this method, we can score recognition (1) *for any number of things presented,* (2) *for any number selected (not necessarily one-half of the number presented),* and (3) *for any number of correct or incorrect things selected.* This may be seen from Table II, which can be extended indefinitely. In scoring, we can make use of this table and save a tremendous amount of time.

One must remember that these values are the most probable values. That is, in the long run, they will measure actual efficiency. Another method, and a more logical one, to score probability is to take, for instance, a performance of X correct choices and 12 — X incorrect ones, and see how probably that performance would happen, supposing the subject actually knows only 1, 2 or any number of the correct things selected. Score the number that the subject is supposed to know in order to make such a performance most probable.

However, this method does not serve our purpose, because it sometimes gives ambiguous results. For example, if a performance is 9 correct to 3 incorrect, we do not know whether the subject should be credited with 8 or 9 which he is supposed to remember. For

Supposing he knows 9, then the probability for him to get

9 correct is	....	220/455
10 " "	....	198/455
11 " "	....	36/455
12 " "	....	1/455

Supposing he knows 8, then the probability for him to get

8 correct is	....	495/1820
9 " "	....	880/1820
10 " "	....	396/1820
11 " "	....	48/1820
12 " "	....	1/1820

But the highest probability of the first column is equal to that of the second column.

$$\frac{220}{455} = \frac{880}{1820}$$

Shall the subject be credited 8 or 9?

TABLE II

Table for Scoring Recognition Memory, 24 presented, 12 correct + 12 incorrect

NO. CORRECT	NO. SELECTED 1	2	3	4	5	6	7	8	9	10	11	12
1	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00
2		09.03	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00
3			18.06	06.57	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00
4				29.54	12.31	05.42	00.00	00.00	00.00	00.00	00.00	00.00
5					39.40	26.78	10.11	04.82	00.00	00.00	00.00	00.00
6						49.05	34.93	25.03	09.03	04.52	00.00	00.00
7							57.82	48.06	32.40	24.07	08.51	04.42
8								66.52	55.39	47.40	31.21	23.76
9									74.96	66.39	53.91	47.17
10										83.32	73.25	66.33
11											91.67	83.31
12												100.00

(5) Reconstruction. In this method, we encountered the same difficulty of chance success, and so far as we know, no attempt was made to eliminate this factor in previous studies. We first divide our problem into (A) position and (B) sequence. It is evident that, by chance, one may put a part of the series in the original position and sequence.

(A) Take position first. Assuming perfect chance in the rearrangement of n things, the number of ways for X of the n things to be out of the original order is the number of n things taken X at a time, minus the number of permutations of X things in which the X things are *not all* out of the original positions. For 12 things, the probability for any number of them to be out of position as compared with any other number, is

0 positions out,	1
2	66
3	440
4	4455
5	34848
6	244860
7	1468368
8	7342335
9	29369120
10	88107426
11	176214840
12	176214841
	479001600

Taking 17621841 as 100% chance and following the same procedure as in the recognition method, we have the results as given in Table III.

TABLE III

POSITIONS OUT	PRELIMINARY SCORE	CHANCE	FINAL SCORE
12	00.00	100.00	00.00
11	09.09	100.00	00.00
10	18.18	50.00	09.09
9	27.27	16.67	22.72
8	36.36	4.17	34.84
7	45.45	.83	45.07
6	54.55	.14	54.47
5	63.64	.02	63.63
4	72.73	.00	72.73
3	81.82	.00	81.82
2	90.91	.00	90.91
0	100.00	.00	100.00

The principle applied in the development of this scale is not limited to a series of any particular number.

(B) Sequence. Here we failed to formulate a simple mathematical statement of the relative amounts of chance and had to depend upon empirical data. By casting a series of 12 members 1000 times and recording the chance sequences of the members, we obtained the results given in Table IV.

TABLE IV

	MEMBERS OUT OF ORIGINAL SEQUENCE											
	1	2	3	4	5	6	7	8	9	10	11	
1st 100 trials	0	0	0	0	0	0	1	5	15	42	37	100
2nd	0	0	0	0	0	1	0	4	13	41	41	100
3d	0	0	0	0	0	0	2	5	10	35	48	100
4th	0	0	0	0	0	0	1	4	19	34	42	100
5th	0	0	0	0	0	0	0	4	14	45	37	100
6th	0	0	0	0	0	0	1	5	20	32	42	100
7th	0	0	0	0	0	0	1	3	18	37	41	100
8th	0	0	0	0	1	0	0	4	12	37	46	100
9th	0	0	0	0	0	1	1	3	18	33	44	100
10th	0	0	0	0	0	0	1	3	13	36	47	100
	0	0	0	0	1	2	8	40	152	372	425	1000

Comparing the totals with each of the 10 groups, we concluded that these results were regular enough to be a valid sample. We then followed exactly the same procedure as in the scoring of position and obtained the final scale as given in Table V.

TABLE V

0 out of sequence,	Score 100.00
1	90.91
2	81.82
3	72.73
4	63.64
5	54.42
6	45.24
7	35.68
8	24.70
9	11.68
10	1.13
11	.00

(C) After scoring both position and sequence, the average of the two was taken as a rough measurement of Reconstruction memory. This final process is arbitrary and may be entirely superfluous. It does not furnish any more adequate measurement of reconstruction than position and sequence taken separately, since the resulting values do not lend themselves to a clearer interpretation.

Quantitative Data

The results from these five methods of measurement are tabulated in Table VI and graphically represented in Fig. II.

TABLE VI

Percentage of Retention

	20 min.	1 hr.	4 hrs.	1 day	2 days
Anticipation*	67.8	50.2	39.0	17.8	10.0
Relearning*	75.0	65.9	54.9	52.1	47.7
Written reproduction	88.1	82.1	60.5	39.2	26.7
Reconstruction	91.5	89.7	75.4	50.9	38.6
Recognition	97.8	94.6	93.3	74.6	71.5

From Table VI and Fig. II, the phenomena of retention may be generally stated as follows:

(1) The amount of retention decreases with time.

(2) On the whole, forgetting is most rapid at first, but there are two notable exceptions. A. The curve for recognition slopes down much more rapidly from 4 hrs. to 1 day than from 1 to 4 hrs. B. In reconstruction, the decrease in the amount of retention is more rapid from 1 hr. to 4 hrs. than from 20 min. to 1 hr.

(3) All the curves are relatively uniform and can be described by mathematical formulae.

*In the latter part of this series of experiments, it was found necessary to counteract the effect of diurnal variations by distributing the 4-hr. series more carefully. Some of these series were given at the regular learning time of the day for each subject; the rest were given either 4 hrs. before or after that time. This procedure was carried out only for anticipation and relearning. Subject K's records were excluded from the above averages for not being so distributed.

Because of a mistake in scoring, subject D's 5 records had to be excluded from the averages for anticipation.

Altogether, the averages for anticipation represent 28 records from 8 subjects; relearning represents 33 records from 9 subjects; the rest include 35 records from 10 subjects.

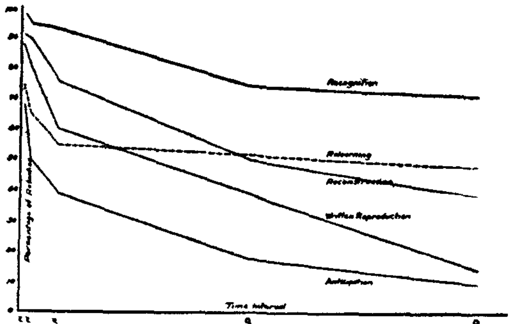


Fig. I

(4) Four of the curves stand invariably in a given order. Recognition gives the highest values. Reconstruction occupies the second position. Written reproduction follows as a close third. Anticipation always has the lowest value.

(5) The relation of the values for relearning varies with the time interval. As may be seen, its order is fourth for the 20-min. interval and second for the 2-day interval.

In general, these conclusions are in harmony with the results of earlier investigations, except those of Ballard. These results are brought together in Table VII and compared with those of the present study.

In relearning, the present results approach most closely to those of Finkenbinder. In no instance is the difference more than 10%. Finkenbinder used practically the same technique as our own, excepting that we required the subject to spell the syllable instead of pronouncing it as a whole.

With the "saving method," an increase in the number of presentations in learning increases the amount of retention, while an increase in the number of presentations in relearning decreases the same. Since the number of presentations in relearning is almost always smaller than that of learning (except in cases of 100% forgetting), every increase of equal or nearly equal magnitude in the number of presentations in both learning and relearning will increase the amount of forgetting, or, in other words, it decreases the amount of retention thus calculated.

TABLE VII

	Relearning					*Recognition*	
INTERVALS	EBBING.	RADOSSAWLJEVITCH	*	FINK.	LUH	STRONG	LUH
15 sec.						84.6	
5 min.		97.5	95.9			72.7	
15 min.						62.7	97.8
20 min.	58.2	88.6	89.8		75.0		
30 min.				75.0		55.5	94.6
1 hr.	44.2	70.7	75.3	72.8	65.9	57.3	94.6
2 hrs.				69.4		47.2	93.3
4 hrs.				64.4	54.9	50.6	93.3
8 hrs.	35.8	47.4	66.5	65.5		40.6	
12 hrs.				63.8		41.1	
16 hrs.				63.0			
24 hrs.	33.7	68.9	70.2	57.8	52.1	28.8	74.6
36 hrs.				58.8			
2 days	27.8	60.9	72.3	55.6	47.7	22.9	71.5
3 days				52.1			
4 days						19.3	
6 days	25.4	49.3	59.5				
7 days						9.6	
14 days		41.0	51.4				
21 days		37.8	48.6				
30 days	21.1	20.2	27.0				
42 days						6.3	
120 days		2.8	3.3				

* Averages reconstructed from R's principal experiments which included only series of 12 syllables. The column immediately to the left, as quoted by F., represents averages from 8, 12 and 16 syllable series.

Other things being equal, it takes a subject more presentations to spell a series correctly than to pronounce it correctly, both in learning and relearning. This may explain the fact that the relearning values reported in the present study are lower than those of Finkenbinder.

Similarity of technique does result in proportional similarity of quantitative data. It points toward the possibility of establishing norms, though not one norm or one general curve of forgetting, for various conditions of retention. Other numerical differences between the present study and previous reports may be easily accounted for by individual differences and the disparity of methods.

The difference between our recognition values and Strong's is sometimes as high as 50%. There is no cause for wonder, however, since the data were collected under as divergent conditions as imaginable. The material, the degree of learning and the methods of scoring were all different. Our data will have to be distorted a great deal before they can conform to the Ebbinghaus type of curve. Even if we could derive a general equation which satisfies both sets of values, it would be so complex and obscure that scientific interpretation would be better off without it. Here lies the danger of speculation without specifying the conditions and variables which enter into the determination of our values.

Miscellaneous Comparisons

1. The Difference between Scoring the Number of Syllables and of Letters in Anticipation.

There are altogether 12 syllables, or 36 letters, in each syllable series. Since a syllable may be only partly anticipated, *i. e.,* when only one or two letters of the syllables are anticipated in the correct position and sequence, it is evident that scoring the number of letters will give higher values than scoring the number of syllables. The increment is, however, very small, as can be seen from Table VIII.

TABLE VIII

	29.min.	1 hr.	2 hrs.	1 day	2 days
Anticipation, scoring 12 syllables	67.8	50.2	39.0	17.8	10.0
" " 36 letters	70.2	54.2	41.6	19.7	10.5

That there is a positive difference at each interval in favor of the number of letters indicates the existence of partial retention which is not ready enough for successful anticipation but which is nevertheless effective. One might anticipate this fact from the nature of the case, quite independent of the magnitude and the probable error of the difference. Such partial retention may be due to one of two causes, or to both. (1) One part of a syllable may be forgotten more rapidly than another part. (2) The association link may be so weakened that it cannot be reinstated within the short time limit of 2 sec. We shall later see that extending the time limit for recall increases the score by a far greater amount than does the present process. A time limit is detrimental to partial retention.

It is important to notice that by scoring the number of letters instead of syllables, the shape of the curve is not materially changed. In Fig. III, the curves for the number of syllables and for the number of letters may be compared by direct inspection.

2. The Difference between Scoring the Whole Series With and Without the First Syllable.

This was done on *a priori* grounds. Apparently, there seems to be a difference between the anticipation of the first syllable of the series and of the other eleven. For each of the eleven, anticipation is facilitated by that part of the series which is already exposed. The associative bond is aroused by so many "cues," such as visual, auditory and vocal, which are not available for the first syllable. Such is not the case with written reproduction. For there the subject may begin from any point of the series and run forwards and backwards.

Specifically speaking, the ground for singling out the first syllable of the series is perhaps insecure. Association is effective not only between immediately consecutive syllables, but also in the most criss-cross way imaginable.[5] So even the recall of the first syllable will be helped by the anticipatory re-instatement of the succeeding members. But in general, one may consider the association between two like members of a series as quantitatively different from that between one of these members and another dissimilar factor.

When the first syllable was thus excluded, we obtained the comparative results of Table IX.

TABLE IX

	20 min.	1 hr.	2 hrs.	1 day	2 days
For 12 syllables..................	67.8	50.2	39.0	17.8	10.0
" 11 "	64.9	46.0	32.0	13.0	8.3

Quite beyond our expectation, the averages for 12 syllables exceed those for 11 at every interval. The first syllable was more

[5] Cf. Ebbinghaus, *Op. cit.,* Ch. IX.

often correctly anticipated than the average of the other 11. Thus the lack of associative "cues" seems to be more than compensated by the favorable effect of primacy, and our *a priori* conclusions become groundless.

3. The Difference between Scoring and Not Scoring Position and Sequence in Written Reproduction.

As already stated, written reproduction was scored (1) as to the gross amount of retention and (2) as to that factor plus position and sequence. In the latter case, the gross amount was scored ½ and position and sequence ¼ each. The comparative results are presented in Table X.

TABLE X

	20 min.	1 hr.	4 hrs.	1 day	2 days
With position and sequence.......	80.3	73.1	50.3	32.5	22.6
Without position and sequence....	88.1	82.1	60.5	39.2	26.7

Scoring position and sequence apparently decreases the amount of retention. It is, of course, much more probable for a syllable to be reproduced than to be reproduced in the original position and sequence. On the other hand, it is also remarkable that there are not so many instances in which a syllable is remembered as to its position and sequence but only vaguely in specific content. The above differences would certainly be more prominent if we had taken into account the extent of chance in our method of scoring.

Later we shall see that the decrease in the amount of retention due to scoring position and sequence is not limited to the conditions of this series of experiments. The phenomenon was reproduced in a second series of experiments in which the degree of the original learning was varied. *Still the difference in amount is not such as to change the general shape of the curve.* Compare Fig. III for this series of experiments.

4. The Difference between Scoring Preliminary and Final Records.

In describing the methods of scoring, we mentioned that records were taken of the amount of error in written reproduction, of the amount of material recalled upon the lapse of preliminary time limits in recognition and written reproduction and of the amount of time spent in the whole process of recall in recognition, reconstruction and written reproduction. These results will be presented and compared in another section.

5. *Corroborative Data from the Second Series of Experiments.*

In the second series of experiments, the Reproduction, Recognition, and Reconstruction methods of testing results were used when the material was learned with different degrees of mastery. When the degree of learning was exactly the same as in the first series of experiments, the two series of values corroborated each other to a remarkable extent. The values for both series of experiments are presented in Table XI. A fuller description of the cond.tions of the second series is to be found at the beginning of Chapter III.

TABLE XI

	20 min.	1 hr.	4 hrs	1 day	2 days
Written reproduction					
1st series of experiments......	88.1	82.1	60 5	39.2	26.7
2d series of experiments.....	90.6	85 8	64.8	45.6	40.2
Average	89 4	84 0	62 6	42 4	33.5
Scoring position and sequence					
1st series of experiments......	80 3	73.1	50.3	32.5	22 6
2d series of experiments......	86.5	81.2	58.0	37.5	33.5
Average	83.4	77.2	54.2	35.0	28 1
Recognition					
1st series of experiments......	97.8	94.6	93 3	74.6	71.5
2d series of experiments......	95.8	95.0	91.6	77.6	78.9
Average	96.8	94.8	92.5	76.1	75.2
Reconstruction					
1st series of experiments......	91.5	89.7	75.4	50.9	38.6
2d series of experiments......	89.3	90.4	74 9	48 6	44.0
Averages	90.4	90.1	75.2	49.7	41 3

For the shorter intervals, the validity of either set of values is self-evident and beyond question. But the difference between the two sets increases with the length of the interval. This fact is to be later considered as a characteristic of individual differences in the ability to recall.

In written reproduction, when position and sequence were scored, the values of the first series were not so closely reproduced in the second series for the short intervals but more closely reproduced for the long intervals. The validity of the long interval values is questionable. As to the short intervals, we have already stated that the method of scoring position and sequence is not very reliable.

Comparison and Interpretation of Data

From Fig. II, two general phenomena are easily observable. (1) The curve for relearning does not fall as rapidly as the other curves and it intersects with the reconstruction and written reproduction curves as the length of the interval is increased. Were it possible to fit each series of empirical data to an ideal curve or family of curves, one would still be confronted with the difficulty that relearning does not satisfy quite the same type of equation as the other memory processes. By increasing the number of constants, we might represent all the series of values by a general logarithmic equation which applies to all conditions, but our ignorance of the actual and specific course of forgetting would be as profound as ever. Suppose that the values of all the constants are given or calculated, which is a wild supposition in the light of our present knowledge of memory processes. We could then be sure of only one thing, viz., *Relearning and the other processes do not satisfy the same type of logarithmic equation.*

(2) With the exception of the relearning curve, the other curves are more or less similar. The similarity becomes more prominent at the end of the 4-hr. interval. It may even be said that after the lapse of that interval, the difference between any two memory processes except relearning and probably recognition, measured at any time, is a definite amount which is constant for those two processes. Recognition is in many respects similar to relearning. It favors the longer intervals and partial retention.

This similarity of the curves is reproduced in the second series of experiments, as may be seen in Table XI.

(1) *Comparison of the Relearning and the Other Curves.*

One possible reason for the disparity as discussed under (1), one might assume, may be traced back to defects in the methods of measurement, which do not take into enough consideration the amount of partial retention. As we shall later develop, the processes of memory fade away gradually, from complete retention to bare recognition. One may thus be led to expect that the amount of partial and uncertain retention increases in direct proportion to the length of the interval. Could we devise a finer method by which each memory process is measured in its *entirety*, the score for each interval would be increased by the amount of partial retention. But the increment would make very little difference in the shape of the curve for the shorter intervals when the total amount of partial retention is rather small; and in fact, we see that the fall of the relearning curve is similar to the others for the short intervals. In other words, the defects in the methods of measurement could not be so easily detected when the interval is short. For the long intervals, these defects could be remedied only by carefully scoring the amount of partial retention. Since the latter is thought to increase with the time interval, the results thus scored would manifestly be a curve which slopes down much less abruptly and approaches a type like that of relearning.

Now the argument as here presented assumes at least two things. A. A more accurate method of measurement will bring to light the amount of partial retention so much so that the shape of the curve will be changed. B. The amount of partial retention increases with the time interval.

As to B, we suggest that this assumption is not always true and shall try to demonstrate the fact in Ch. IV.

Assumption A seems more plausible only because we cannot develop a method of such magical accuracy. If the methods of measurement were the only faulty factor, then the improvement of technique by way of refining these methods would in proportion make the other curves approach more closely the relearning type of curve. However, within the limits of our investigation, the shape of the curves do not materially change on account of minute variations in the methods of measurement. The following facts indicate that the differences under consideration are more fundamental than merely a matter of technique.

(a) In anticipation, scoring the number of letters as well as syllables increases the value by including partial retention, *but the increments are not such as to make the curve fall less rapidly than does scoring the number of syllables alone* (Fig. III).

(b) In written reproduction, scoring position and sequence in addition to the amount of reproduced material does not change the shape of the curve very appreciably. When position and sequence are not scored, more allowance is made for partial retention, for a syllable which is retained only as to its content and not its relative order is scored as much as one which is completely retained. According to the proposed theory, scoring position and sequence would make the curve fall much more rapidly. This is not a fact.

(c) In recognition and written reproduction, as will be discussed later on, extending the time limit for recall so as to make more room for the reinstatement of partial retention does not change the general shape of the curve, though the increment of the score after the extension of the time limit in recognition does increase with the length of the interval so that the curve slopes toward the abscissa more gradually. (For recognition in the first series of experiments, see Fig. III.)

Thus, scoring partial anticipation, neglecting position and sequence in written reproduction and, finally, extending the time limit for recall all fail to eliminate the apparent difference between the relearning curve and the other curves. The assumptions of the theory cannot be substantiated.

It is doubtful whether the relearning curve can be directly compared with any of the rest. Relearning, being a composite method, may be analyzed into anticipation and subsequent learning. Most probably, the subsequent learning is a function of the amount of anticipation. The more the amount of anticipation, the less the number of presentations necessary for relearning.

If any such causal connection be found to be generally valid, then the relearning values can be constructed from the original learning and the amount of anticipation. Only such analysis could bring out any natural relationship that may exist between the

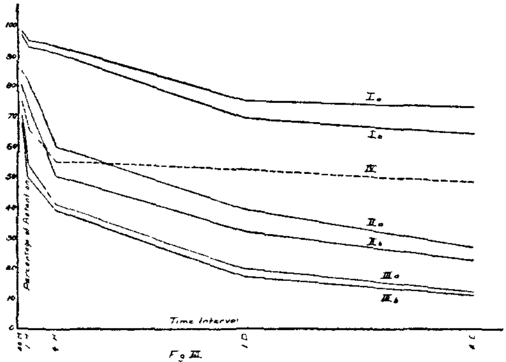


Fig. III.

amount of retention and the time necessary for relearning. For the present these two factors cannot but be treated as independent values, particularly for the following reason.

The values for relearning are based upon the time of learning and relearning, the proficiency of retention being measured by a ratio of time; while the values for the other curves are built on an almost totally different criterion. With the latter, the standard of proficiency is not the length of time, but the ability to retain the whole series, and different degrees of retention are measured as steps approaching that standard.

For this reason, the time for relearning cannot be interpreted in terms of the amount of retention, but both must be taken into consideration in order to understand the phenomenon of retention quantitatively. A score made by the "saving method" cannot be converted into another score except as to mean the amount of time saved. As to the measurements for the amount of retained material, the same course of reasoning will lead to the conclusion that these measurements do not indicate anything as to the relative difficulty experienced in mastering the material. For instance, 60% of retention is three times as high as 20% with respect only to the actual amount of material; it explains nothing as to how it was acquired, or how much it is really worth as compared with another amount acquired with the expenditure of a larger or smaller amount of time.

Perhaps a consideration of the shape of the learning curves for memory as involved in the relearning method will bring to mind more distinctly the disparity between the two standards of measurement, namely, time and amount. If in both learning and relearning, the effect of each presentation of the series upon learning were constant, the learning curves would then satisfy equations of the first degree such as hypothetically represented in Fig. IV. Further, if the effect were constant for relearning

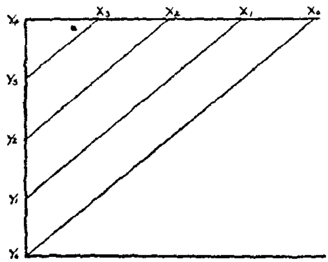


Fig IV

after any time interval and with any amount of actually retained and reproduced material, then the curves would be parallel. In such a case, the amount of retention would be a *function of the* time necessary for relearning, and *vice versa*. For, referring to Fig. IV, the amounts of retained material after different time intervals are Y_0Y_3, Y_0Y_2, etc. Correspondingly, according to the "saving method," the percentages of retention and forgetting are calculated according to distances on the line parallel to the abscissa which represents the number of presentations necessary for learning or relearning. X_0Y_4 is the number of presentations in learning. X_3Y_4 is the number of presentations necessary for relearning after the lapse of an interval when the amount of retained material as represented on the ordinate, is Y_0Y_3. The percentage of forgetting for that interval is $X_3\ Y_4\ /\ X_0\ Y_4$. It corresponds to the amount $Y_3\ Y_4$. Similarly, $X_2\ Y_4\ /\ X_0\ Y_4$ corresponds to $Y_2\ Y_4$, $X_1\ Y_4\ /\ X_0\ Y_4$ corresponds to $Y_1\ Y_4$, etc. It is clear that since the learning curves are supposed to be parallel, the percentages of forgetting, or of retention, thus calculated from the abscissa would always be proportional to the percentages calculated from the ordinate. *Then the curves of retention based upon the "saving method" and upon the amount of recalled material would be similar, when plotted on the same scale.*

As a matter of fact, the learning curves for memory are not parallel straight lines. Practically all the curves so far determined are negatively accelerative, including especially the recent work of Kjerstad.[6] This is also corroborated by our own results in both learning and relearning. The degree of negative acceleration in the relearning process also varies with the time interval or the amount of actually retained material. That is, if the curves follow the same law of negative acceleration, they can then be represented as in Fig. V. In the figure, Y_3 Y_4, Y_2 Y_4, etc. represent respectively the amounts of forgetting after different time intervals, X_3 Y_4, X_2 Y_4, etc. represent respectively the corresponding number of presentations necessary for relearning for each of the assigned time intervals. From that figure, two facts become self-evident.

a) Not only do X and Y differ in the scale and unit of measurement, but the functional relationship between the two is not so simple as that represented in Fig. IV and not such as could be easily determined. Referring back to Fig. II, this difference in the units of the scales and this complex and unknown functional relationship between them would mean that the absolute height of the retention curve for relearning, as compared with that of the other retention curves, cannot be interpreted by mere inspection. The absolute values of the curves cannot be directly compared.

b) But the more significant fact is that these very characteristics of the learning curves will directly lead to the particular difference between the shape of the relearning curve for retention and of the other retention curves, as seen in Fig. II. In Fig. V, the number of presentations, X_3Y_4, X_2 Y_4, etc. increases *at first* more rapidly than the corresponding amounts of forgetting,

[6] *Op. cit.*

Y_3, Y_4, Y_2, Y_1, etc.; but less rapidly when the amount of actually retained material is small, *i. e., when the time interval is increased.* Now if we plot the values on the ordinate against the variable of time in such manner as manifested by the retention curves

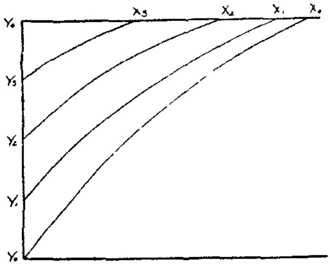


Fig V

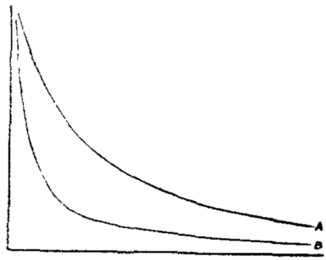


Fig VI

for the amount of reproduced material (A), and if we further draw a curve (B) for the retention values as determined from the abscissa on the basis of (A), their relationship will be such as given in Fig. VI. The difference between the two is exactly what we observed in Fig. II.

Thus, the difference between the relearning curve for retention and the other retention curves is not due to the phenomenon of retention as such, but to the characteristic of the learning curve. If the latter were invariably a simple logarithmic curve, it would necessarily follow that the *forgetting* curve as determined by the "saving method" would be logarithmic. When this relationship obtains, as claimed by most of our predecessors, it is important to remember that,

a. The curve of forgetting for the saving method is logarithmic only because it involves the ratio of two logarithmic learning curves.

b. Similar phenomena cannot be expected to reappear when the same factors are not involved.

The Unsatisfactoriness of the Relearning Method.

It is now evident that neither time nor amount is a complete measurement for retention, but both are not equally *convenient.*

A. The first unsatisfactoriness of the relearning method is that it achieves much less validity of data for the same expenditure of time as compared with the other methods.

As a test for validity, we use the mean deviation of the values of all the series which a subject learns or reproduces for each specific interval. This M. D. may be made directly comparable with other M. D.'s by applying Pearson's formula $V = \frac{100 \text{ M. D.}}{\text{Median}}$. The greater the coefficient V, the less is the validity. Five of the subjects learned five series for each of the five intervals in the first series of experiments. Five is a ridiculously small number for such statistical treatment. If, however, all the results point uniformly in one direction, most probably there is an actual difference in the degree of validity of these values. Table XII presents the results for only two intervals, 20 min. and 1 hr. Beyond that, the variability of the relearning method becomes so large as to make such comparison superfluous and meaningless. A ? mark indicates that, of the two intervals and the two particular methods compared, the coefficient of variability is larger for one method at one interval, and for the other at the other. In a similar way, a + sign means larger variability for relearning at both intervals.

TABLE XII

SUBJECT	ANTICIPATION	RECOGNITION	RECONSTRUCTION	W. REP.
D	?	+	—	?
Le	?	+	?	+
Lo	?	—	+	?
R	+	?	?	?
W	?	+	+	+

Of all the paired comparisons, none is distinctly in favor of the relearning method. Its coefficient of variability is higher than that of any other method. That is, for equal expenditure of time and energy, relearning produces less satisfactory results than does any other method.

B In addition to the above, the relearning method makes impossible certain correlation studies which are easily accessible to the other methods. Questions like the following have to be answered one way or another. (a) What is the relation between the speed of learning and the amount of retention for each interval? (b) For all the series learned by the same subject, what is the relation between the difficulty of learning and the amount of retention, etc.? As to question (a), the problem of the correlation between "immediate" and "permanent" memory has claimed a goodly number of working days. Very little has been written on question (b), but Ebert and Meumann[7] once raised the problem whether subject matter that is readily learned is forgotten more rapidly or more slowly than that which requires greater labor in memorizing. In dealing with these problems, the relearning method was generally used. Relearning, as we already discussed, involves two learning processes. Suppose a number of syllable series to be learned in X, X′, X″, . . . presentations and relearned in Y, Y′, Y″, . . . presentations. The amounts of forgetting will be $\frac{Y}{X}, \frac{Y'}{X'}, \frac{Y''}{X''}, \ldots$ Learning and forgetting involve the same factor X, X′, X″, . . . That fact deprives any correlation study of its real significance. If it is further argued that the percentage of retention is determined by X—Y, X′—Y′, X″—Y″, ... which are independent values, we suggest that these values can be more reasonably established by some other methods which directly measures the amount of reproduced material.

(2) *Comparison of the Curves which Represent Only the Amount of Reproduced Material.*

We have explained why the relearning and the other curves are different in general form, and now proceed to the problem why these other curves differ in absolute numerical value. But we cannot treat the problem without first proving that the numerical differences are real and not adventitious.

First, we can refer back to the data for each individual subject, which constitute the group averages. From the group curves, recognition has a higher value than reconstruction for each of the five intervals. The same relation holds for reconstruction and written reproduction, and for the latter and antici-

[7] "Ueber einige Grundfragen der Psychologie der Uebungsphenomene im Bereiche des Gedachtnisses," *Archive f. Gesamte Psy.*, Vol. IV, 1904, pp. 193-4.

pation. This particular order of proficiency for the several measurements of retention holds true for the majority of the subjects. Thus, of the 20 individual values, 10 from each series of experiments, which constitute the final recognition score for the 20 min. interval, only one is higher than its corresponding value which goes to make the final score for reconstruction for the same interval, etc. Table XIII presents these facts in brief.

TABLE XIII

PERCENTAGE OF INDIVIDUAL CASES WHICH CORRESPOND TO GROUP RESULTS

	20 min.	1 hr.	4 hrs	1 day	2 days
Anticipation cf. written reproduction*	100	100	87	87	75
Written reproduction cf. reconstruction	60	75	75	70	65
Reconstruction cf. recognition	95	80	95	95	95

* Only 8 values from 8 subjects in the 1st series of experiments. The rest are constituted of 20 values from 20 subjects in both series of experiments.

Secondly, the validity of the numerical differences may be proved by the magnitude of their probable errors, as presented in Table XIV.

TABLE XIV

	20 min.	1 hr.	4 hrs.	1 day	2 days
Difference between					
Antic. and Wr. Rep............	20.3	31.9	21.5	21.4	16.7
P. E. of difference..............	4.5	3.7	5.5	7.0	6.1
Difference between					
Wr. Rep. and Reconstruction...	1.0	6.1	12.6	7.3	7.8
P. E. of difference	1.1	1.7	2.7	4.5	4.3
Difference between					
Recons. and Recognition......	6.4	4.7	17.3	26.4	33.9
P. E. of difference..............	1.5	1.9	2.2	4.3	4.3

The P. E. of the difference between reconstruction and written reproduction can be reduced by increasing the number of subjects. In other words, the values for reconstruction and written reproduction may be slightly changed with the increase in the number of subjects, especially for the 20 min. interval. That there is a positive difference in favor of reconstruction at each of the intervals is unmistakable. The validity of all the other differences is self-evident.

Returning now to the group curves, we further find that, while after the lapse of the 4 hr. interval the course of the curves become more or less parallel, such is not the case for the shorter intervals. There the decreases in value are more abrupt for certain curves than for others. On the whole, the one that begins the lowest at 20 min. falls the most rapidly so that the curves become farther and farther apart as the length of the interval increases. So the Ebbinghaus tradition is substantiated by our data to that extent.

However, the negative acceleration theory of the curve of forgetting does not hold for all of our values. The two notable exceptions, as may be seen from pp. 21-22 are:

1. In recognition, the transition from 1 hr. to 4 hrs. is much less accelerative than that between 4 hrs and 1 day.

2. In written reproduction, the curve approaches a straight line between 20 min. and 4 hrs.

A significant fact is that these exceptions in the first series of experiments were reproduced in the second series. Further, imposing a time limit upon the act of recognition and scoring position and sequence in written reproduction did not in the least ameliorate these peculiarities. Is the curve of forgetting a logarithmic curve, as has been repeatedly maintained? These facts must be taken into consideration and explained.

One must also remember the course of the curves after the lapse of the 4-hr. interval, as described in a preceding paragraph. What is its bearing on a general logarithmic equation?

From the above, we may still conclude that, of two memory processes, the one that commands a higher "initial" amount of forgetting tends to fade away more rapidly than the other. This acceleration approaches a limit at the end of 4 hrs., and does not resume its initial course even to the end of 2 days.

One word as to what is meant by an "initial" amount of forgetting, a term first used by Bean. When retention is measured at the end of 20 min., we see only a cross-section of the stream of forgetting, to use James' old metaphor. From our data we cannot determine whether the grade of the headwaters is steep or level, or where the grade actually begins. In other words, we do not know what the "initial" amount of forgetting is, for each process. The quantitative differences as we measure at the end of 20 min. may doubtless be traced back to differences of the same kind. We assume that, if at 20 min., retention as measured by the anticipation method is less than by the recognition method, this was also true for shorter periods than 20 min., though not in the same ratio and by the same amount. But in fact, forgetting for one process may not begin at the same point as for another.

This is particularly true for written reproduction, recognition and reconstruction. When the method of presentation in the original learning is different from the method of testing, we can no longer compare directly the amount of immediate retention with the values that determine the curves of forgetting. In anticipation and relearning we know that the amount of retention immediately after learning is 100% so that we can trace these curves back to Y_0, *i. e.*, where the length of the interval is 0. But in the other three methods, the forgetting curves begin at 20 min. and we cannot go back any further. It is perhaps unfortunate that we did not vary the methods of the original learning as well as the methods of testing, but then we would have introduced another constant and would have made our data more difficult to interpret.

Very likely, the quantitative differences between any two processes measured at any time may be greatly reduced by transposing the curves so that the points of "initial" forgetting coincide. All this may be empirically determined. Unless such problems are solved, the term "initial" forgetting will remain as meaningless as the "logarithmic" curve of forgetting.

Possible Explanations for the Quantitative Differences.

The conditions are complicated. Perhaps no single explanation is sufficient, but the following are more than likely.

1) The temporal order in which the methods were applied. In one half of the series given to each subject, anticipation was tested before relearning. In the other half, written reproduction was the first test applied. Recognition followed after written reproduction, and reconstruction after recognition. The high

retention values for recognition and reconstruction and the low values for anticipation and written reproduction may partly be due to the presence or absence of a preceding recalling process. For the present, we cannot determine whether that effect actually exists and to what extent. It is probable, however, that the effect of a preceding recalling process may simply reduce the duration of the succeeding recalling process without changing the amount of recall of the latter. In Ch. VI we shall see that the average duration of the written reproduction process was longer than for recognition, and the latter longer than for reconstruction. It is possible, of course, that a preceding recalling process affects both the amount and the duration of a succeeding one.

Apparently, the explanation does not cover all the facts. Reconstruction followed after recognition but gave much the smaller values.

2) The duration of the several processes. The duration of the anticipation process was 24 sec.; that of written reproduction, 5 min.; while recognition and reconstruction did not have time limits. So the order of the numerical values corresponds to the order of the duration of the processes. But as we shall later discuss in full, it took the average subject much less than 5 min. to complete the written reproduction process. The actual duration of recognition was not half as long as written reproduction, and that of reconstruction was still shorter. The actual differences in duration are not proportional to the numerical differences in the retention values. Written reproduction had the longest duration but gave the lowest retention values except anticipation.

3) Differences in the units of measurement. In anticipation and written reproduction, the scores were based upon the amount of reproduced material. The requirement of the recognition method was merely that the different members be re-instated upon the presentation of the original series. In reconstruction, only position and sequence were required. The implication would be that the numerical differences under discussion cannot be taken too seriously.

However, this kind of explanation is, in a way, begging the question. In written reproduction, recognition and reconstruction, we scored *the same records* for different values. Were the scales and units identical, we could not have had more than one value. The required explanation is this: Given different methods of scoring the same records, why do the results differ in such a characteristic way? The proposed explanation only tells us that we should not have used those methods. So we are led to the fourth probable explanation which seems to us to be most reasonable.

4) The conditions of recall. The experimental situations under which the subject was required to recall were vastly different. Our measurements took into consideration, among other things,

- a) The retention of the separate parts of a serial act of memory as called for by
  - (a) Written reproduction, and
  - (b) Recognition.
- b) The associative links that connect and combine these parts.
- c) The readiness with which they are recalled, as measured by
  - (a) The amount of time necessary for recall and
  - (b) The amount of material recalled within a given time limit.

These factors are not equally important in all the methods of measurement. *The number of these factors involved and the extent to which they are involved determine the quantitative differences.* Thus, by order of the difficulty of recall, (Figs. II and III), the four methods are ranked:

1. Anticipation.
2. Written reproduction, scoring position and sequence.
2a Written reproduction, not scoring position and sequence.
3. Reconstruction.
4. Recognition, with time limit.
4a Recognition, without time limit.

Analyzing the number of factors involved, they follow:

1. Anticipation. Factors a) and c) are both important. b) is necessary for complete re-instatement, but once a mistake is made, it is automatically corrected by the exposure of the syllable.
2. Written reproduction, scoring position and sequence. Factors a) and b) are equally important. c) is involved only to the extent of being able to reproduce the material within 5 min.

2a Written reproduction, not scoring position and sequence. The conditions are the same as in 2, minus b).

3. Reconstruction. Only b) is involved, being the complementary of 2a.

4. Recognition, with the time limit. a) is only slightly involved, b) not at all and c) only to the extent of being able to select the 12 syllables within 90 sec.

4a Recognition, without time limit. Only a) is partly involved.

From this we conclude that

1. The quantity of recall depends upon the number of restricting factors in the recall situation. The greater number of such factors and the more exactingly they operate, the less the amount of recall.

2. The proportionality between the number of such factors and the amount of recall is an index to the practical validity of our methods of measurement. The scores given by these methods are directly comparable.

It is, therefore, meaningless to say that forgetting in general follows a certain equation. There can be as many curves of forgetting as there are situations and methods of measurement. We know almost nothing as to the "initial" amount of forgetting and very little as to the general shape of any curve. It may be harmless to say that forgetting is a logarithmic function of time, so long as we remember that the significant thing that should influence future investigations is not the logarithm, but the determination of the constants. The latter we cannot deduce from generalizations, but can only measure under variable conditions.

III. RETENTION AS A FUNCTION OF THE DEGREE OF LEARNING

In the second series of experiments, the condition that was varied was the degree of the original learning. The ten subjects of the second group served throughout these experiments, each learning at least 44 syllable series, besides preliminary practice trials. The series and the degrees of learning were distributed according to the following scheme.

1. Altogether we used four degrees of learning.
   - A. 100% learning, with the same conditions as in the first series of experiments.
   - B. 150% learning, in which the subject was given one half of the number of presentations in addition to what was required for the first errorless anticipation of a series. Thus, if a series was learned in 10 presentations, 5 more were given. If 9, also 5.
   - C. 67% learning. The average number of presentations was calculated for each subject after he had learned 20 series besides the preliminary trials. In 67% learning, he was given two-thirds of that number of presentations. (The average number was previously taken to be the total number minus one, *i.e.*, the number of presentations in which there was actual anticipation. In the present case, the *total* number was used.)
   - D. 33% learning. By the same process of computation, one-third of the total number of presentations was given.

2. In 100%, 67% and 33% learning, the intervals used were the same five as in the first series of experiments. Two of these were omitted for 150% learning, but four others were added. So there were seven intervals for 150% learning, viz., 2 hrs., 3 hrs., 4 hrs., 6 hrs., 12 hrs., 1 day and 2 days. The 2-hr. and 3-hr. intervals were arbitrarily chosen after they had been tried out

on several subjects. The object was to select such intervals as would facilitate the comparison of the retention curves for 150% learning with the other curves from the same series of experiments.

Two series were given each subject for each interval and each degree of learning. All the series for 100% and 150% learning except the 2-hr. and 3-hr. ones were completed by each subject before he tried 67% or 33% learning. The average number of presentations for learning these 20 series furnished the required basis of computation for determining 67% and 33% learning. The 2-hr. and 3-hr. intervals for 150% learning were added at the end of the whole series of experiments in order to trace the curves of forgetting further back toward the ordinate. One extra 6-hr. series was given each subject also toward the end of the experiments in order to counteract the effect of diurnal variations.

For each degree of learning, care was taken to distribute the long and short intervals evenly.

Only the written reproduction, recognition and reconstruction methods of measurement were used in recall.

Quantitative Data

The results of the experiments are tabulated in Table XV and graphically presented in Figs. VII-IX.

TABLE XV

	20 min.	1 hr.	2 hrs.	3 hrs.	4 hrs.	6 hrs.	12 hrs.	1 d.	2 d.
Wr. Rep.									
150% learning..			88.0	84.4	81.9	65.6	54.4	38.5	30.8
100%	90.6	85.8			64.8			45.6	40.2
67%	85.4	72.5			65.8			41.5	24.8
33%	67.7	54.0			42.7			26.2	13.7
Recognition									
150%			97.5	95.8	93.3	91.6	92.5	83.2	72.8
100%	95.8	95.0			91.6			77.6	78.9
67%	93.2	93.3			84.7			73.7	61.5
33%	73.3	64.4			54.6			45.7	25.5
Reconstruction									
150%			87.5	92.1	90.8	78.9	81.3	43.4	43.9
100%	89.3	90.4			74.9			48.6	44.0
67%	92.0	77.9			65.3			56.6	31.8
33%	75.6	61.9			48.1			26.0	20.0

1. The Range of Differences.

As may be seen from Table XV, recognition, as a rule, gives the highest value for each interval and for each degree of learning. Reconstruction generally occupies the second place and written reproduction the last. Now if we take the difference between the highest and the lowest values for each interval and for each degree of learning, we can make a comparative study of the range of the differences. The facts are presented in Table XVI.

TABLE XVI

RANGE OF DIFFERENCES

	20 min.	8 hrs.	4 hrs.	1 day	2 days
33%	7.9	10.4	11.9	19.5	12.8*
67%	7.8	20.8	19.4	17.1*	36.7
100%	6.5*	9.2*	26.8	28.0	38.7
150%			11.4*	44.7	42.0

In this table the highest value for each interval is underlined and the lowest marked with an asterisk. The tendency is for the highest value, *i. e.*, the greatest range, to occur at the shorter intervals for the lower degrees of learning and at the longer intervals for the higher degrees of learning. The tendency for the occurence of the smallest range is *vice versa.* These facts reflect the characteristic way in which the different retention curves approach the x—axis with the increase in the length of the time interval.

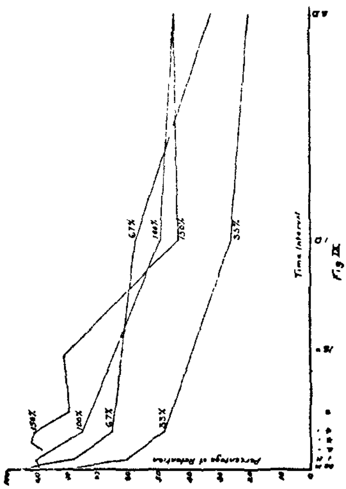
In general, the above range of the differences increases with the time interval, but at different rates for different degrees of learning. Theoretically, a higher degree of the original learning, of course, increases the amount of retention for all the methods of measurement. But it particularly favors the more difficult methods for the shorter intervals so that the range of the differences between the easiest and the most difficult methods is small. As the time interval is lengthened, this advantage rapidly disappears. So the different curves fall gradually apart, and the range of the differences increases accordingly.

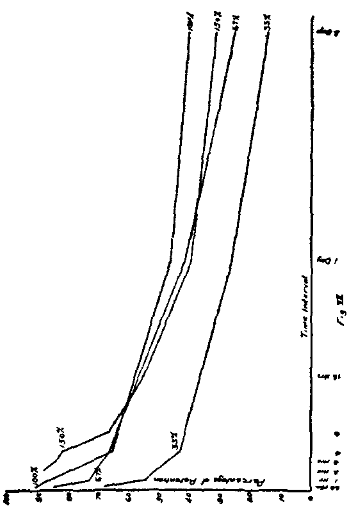
With a lower degree of learning, the effect is generally to decrease the amount of retention for all the intervals and all the methods of measurement, but the special advantage is on the side of the easier, not the more difficult, methods. This effect increases the range of the differences for the shorter intervals, or at least keeps it as large as for the higher degrees of learning, which means that the range will be proportionally greater. Another characteristic result of a lower degree of learning is that, while the curves for all the methods of measurement begin rather low, they fall very slowly and keep almost parallel to each other. The range of the differences is thus kept within a small variation. For the higher aegrees of learning, the curves fall at such different rates that they grow farther and farther apart.

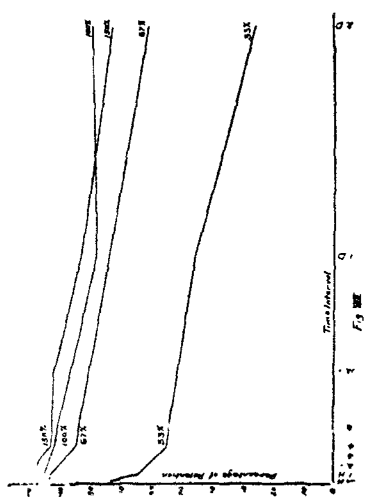
We have, therefore, at least three types of curves resulting from varying the degree of the original learning and at the same time using three methods of measurement. Type 1 begins high and falls slowly. Type 2 begins high and falls rapidly. Type 3 begins low and falls slowly. In order to establish a general formula for all these types, the numerical differences must be more accurately determined.

2. Increase in the Degree of Learning and Diminished Returns in the Amount of Retention.

In Figs. VII-IX one can easily observe that the difference between the curves for 33% and 67% learning is the greatest, that between the curves for 67% and 100% learning very much less, and that between the curves for 100% and 150% learning the least of all. The curves for 100% and 150% learning often cross each other so that it is sometimes difficult to tell whether an increase of 50% of learning actually resulted in any increase in the amount of retention. In written reproduction they cross for the first time at a point whose abscissa represents an 8-hr. interval. Previous to that point the difference between the curves is distinct. There is a more decided difference between the 100% and 150% curves for recognition and reconstruction. Table XVII presents the difference for all three methods of measurement and for all the intervals. The probable errors of the differ-







ences are not included. They are very high for the smaller differences. The larger differences are self-evident.

TABLE XVII

		20 min.	1 hr.	4 hrs.	1 day	2 days
Written Reproduction						
Difference	33 and 67%....	17.7	18.5	23.1	15.3	11.1
	67 and 100%....	5.2	13.3	—1.0	4.1	15.4
	100 and 150%....			17.1	—7.1	—9.4
Recognition						
Difference	33 and 67%....	19.9	28.9	30.1	28.0	36.0
	67 and 100%....	2.6	1.7	6.9	3.9	17.4
	100 and 150%....			1.7	5.6	—6.1
Reconstruction						
Difference	33 and 67%....	16.4	16.0	17.2	30.6	11.8
	67 and 100%....	—2.7	12.5	9.6	—8.0	12.2
	100 and 150%....			15.9	—5.2	—0.1

— sign indicates that, of the two values compared, the one for the higher degree of learning is numerically smaller.

At the outset, one might assume that this phenomenon of diminished returns could be due either to (1) practice effect or to (2) difficulties in the original learning.

(1) As the series of experiments occupied more than three months, the proficiency of learning for most of the subjects was somewhat improved. The extent of this improvement we shall develop in another section. After five weeks of practice in 100% and 150% learning, there is no wonder that they could now use the 67% of the average number of presentations to greater advantage. The average number was calculated from the learning of the first 20 series when the practice effect was still increasing. By this factor one might partly explain why the difference between the curves for 100% and 67% learning is so meagre.

But one cannot explain the still smaller difference between the curves for 150% and 100% learning by the same theory. Further, what is true of 67% learning is to a less extent also true of 33% learning. The difference between the latter and all the other curves is the greatest of all. How could the assumed practice effect have influenced 33% learning so differently?

(2) When the degree of learning is reduced to 33%, the

amount of material originally learned is greatly decreased. One cannot be expected to retain what he never learned. So instead of giving the amount of retention for 33% learning as a percentage of the whole syllable series, one might also argue that it should be stated as a percentage of the number of syllables actually learned.

The last statement, however, amounts to saying that instead of 33% or 67% learning, we should have had 100% learning. Doubtless we would then expect the resulting retention curve to coincide with the ordinary retention curve for 100% learning.

In fact, the number of syllables actually reported is a very poor measurement of the degree or the amount of learning. The subject was required to learn the series by anticipation, spelling the syllables aloud. What he did not correctly report after 33% or 67% learning could very often be correctly reproduced even at the end of a comparatively long interval. This phenomenon was particularly manifest for the more efficient learners who required only from 1 to 3 presentations for 33% learning. The effect of the last presentation in the original learning was not brought out by subsequent anticipation. When the total number of presentations was not more than 1 or 2, this last effect became increasingly important. Our records prove that the ability to recall without original correct anticipation occurred with more or less frequency for all the subjects. We may further mention that one of the subjects could not develop the habit of spelling aloud what he learned in the first few presentations.

A far more plausible explanation for this phenomenon of diminished return is that *the effect of the different degrees of learning upon the amount of retention follows the same sequence as does the learning curve for memory.* If negative acceleration is characteristic of the acquisition of immediate memory, as has been repeatedly proved, one may assume that the same phenomenon will reappear in the measurement of permanent memory, which differs from the former only by the introduction of a longer time interval.

From the shape of the learning curves for memory, it has been concluded that the phenomenon of diminished returns holds (1) for immediate recall and (2) for all degrees of learning up to 100%. We can now state in the light of our data that, within the limits of our experiments, this phenomenon also holds (1) for delayed recall and (2) for more than 100% learning.

However, even on this theory one can hardly explain the negative differences in the amounts of retention between 150% and 100% learning for certain intervals. In Table XV, the value for 150% learning is lower than for 100% learning at the end of 1 and 2 days for written reproduction; at the end of 2 days for recognition; at the end of 1 day for reconstruction. At the end of 2 days for reconstruction the values for the two curves are practically equal. It is impermissible to extend the law of negative acceleration to cover these negative cases. One can hardly conceive of an increase of 50% of learning as resulting in an actual decrease in the amount of retention. In Chs. IV and V we shall find that the characteristic deviations for 150% learning are not limited to these unexpected changes in the amount. There we shall offer a general explanation for all these facts.

With the exception of these peculiarities for 150% learning, the phenomenon of diminished returns seems to be quite as general for delayed recall as for immediate memory. Naturally one would look for a common cause for both phenomena unless there are reasons for the contrary. We further maintain that the similarity between these effects is not a coincidence, but almost a mathematical necessity. Given the effect of diminished returns in the curve for immediate retention, one will have to make some wonderful assumptions for not expecting the same effect to appear in delayed recall. This may be clearly seen in Fig. X. The curves presented therein are hypothetical.

When the phenomenon of diminished returns occurs with immediate memory, a simple way to state this fact is to give the general equation of the curve, $y = a - e^{-x}$. When nothing is learned, the curve passed through the point of the origin. Hence,

(1) $y = 1 - e^{-x}$.

The limit of y will be 1, *i. e.*, the mastery of 100% of material regardless of the degree of over-learning.

Supposing that the phenomenon of diminished returns did not reappear in delayed recall, the curve for the retention values corresponding to various degrees of learning would have to be

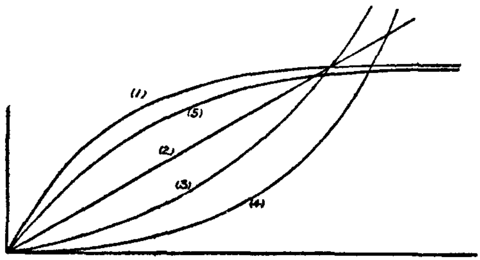


Fig I

either a straight line or one with positive acceleration. For the former, we have the equation $y = mx + b$. When nothing is learned, nothing is retained. Therefore,

(2) $y = mx$.

When the value of m is properly chosen, the curves will intersect with each other as in the figure.

If instead of linear regression, there were more or less positive acceleration, we could also generalize the fact by stating the equation.

(3) $y = Ax + Bx^2 + Cx^3 + \ldots\ldots$

For certain values of the constants the curves (1), (2) and (3) will meet at the same point, which represents that for practical purposes the material is so well mastered that there can be no further forgetting.

Now if we compare the curves between this point of intersection and the point of origin by mere inspection, it becomes clear that for whatever values of the constants, the maximum difference between (1) and either (2) or (3) will not occur in the immediate neighborhood of either of the points, but somewhere in the middle.

Further, curve (1) is fixed. (2) is also fixed when the value of m is determined. Equation (3) is an endless series. x is positive. Assuming all the constants to be positive, then the fewer terms we take, the more abrupt will be the change in curvature, such as curve (4) in the figure.

From this description of the curves one can see that, given the effect of diminished returns in the learning curve for memory, the same effect will not appear in delayed recall only on one of these two conditions.

A. A certain medium degree of learning is the least effective with respect to the amount of retention. This amount will be increased with more learning and also with less learning. Or

B. The degree of learning has no considerable effect upon the amount of retention. It remains very meagre for all degrees of learning until the latter attains a critical value when all of a sudden the amount increases to nearly 100% and forgetting disappears.

Unless and until either one of the conditions is empirically fulfilled, the conclusion still holds true that the phenomenon of diminished returns is general for both immediate and delayed recall. The learning curve and the curve constructed on the basis of the diminishing amounts of retention due to different degrees of learning are related to each other somewhat as (1) and (5). In the latter, the parameter assumes the value a, less than unity.

Ebbinghaus[1] long ago discovered this tendency of diminished returns, though this is not clearly stated in his monograph. Thus he found that the amount of retention after one day was a function of the number of presentations used the previous day. That number was varied from 8 to 64, by intervals of 8. "For each three additional repetitions which I spent on a given day on the study of the series, I saved, in learning this series 24 hours later, on the average, approximately one repetition; and, within the limits stated, it did not matter how many repetitions altogether were spent on the memorization of the series." In Sec. 34, where he treated retention as a function of repeated learning, he concluded, "The effect of the repetitions is at first approximately constant, the saving which results from these repetitions increases accordingly for a while proportional to their number. Gradually,

[1] *Op. cit.*, Ch. VI, Sec. 22-23; Ch. VIII, Sec. 34.

the effect becomes less; and finally, when the series has become so firmly fixed that it can be repeated almost spontaneously after 24 hours, the effect is shown to be decidedly less."

The last conclusion quoted above is self-evident, and clearly corroborates our results. The other statements may be misleading. When the effect of the increase in the degree of learning upon the amount of retention is said to be "constant," or to increase by arithmetical progression, it amounts to saying that, *according to his "saving method," it is negatively accelerative.*

The formula for the "saving method" is $Q = \frac{100\,(L - WL)}{L}$, in which Q is the percentage of saving, L the time required for learning, and WL the same for relearning. The formula holds when L is equal to or greater than WL. When such is the case, a constant numerical increment to both the numerator and the denominator such as in his experiment, $\frac{(L+3) - (WL-1)}{L+3}$, $\frac{(L+6) - (WL-2)}{L+6}$, $\frac{(L+9) - (WL-3)}{L+9}, \ldots\ldots,$ will make each term increase in value, but its difference from the immediately preceding term decrease in value. The effect is thus negatively accelerative.

When L is smaller than WL, as in incomplete learning, Ebbinghaus used as a basis for computation the hypothetical L that would have been spent had it not been for the previous incomplete learning. The amount of saving was a percentage of that hypothetical L, thus neglecting the amount of time that was actually spent on the previous day. If this amount of time had been taken into consideration, then the effect of the increase in the degree of the original learning upon the amount of retention could be shown to be negatively accelerative throughout his investigation.

3. A Further Word Regarding the General Shape of the Curves.

In the last chapter the difficulties for stating a general equation which would satisfy all the phenomena of forgetting were fully elaborated. The reader is now referred to Figs. VII to IX. A general "logarithmic" equation must make allowance for the crossing and recrossing of the "families" of written reproduction and reconstruction curves. Some of these irregularities may indeed be traced back to inaccuracies in the data and may not reappear in another series of experiments of the same kind, but we maintain that any mathematical statement of the Ebbinghaus tradition will require more experimental background.

Perhaps the most embarrassing group of curves is that for recognition. It would be just as easy to fit the corresponding curves of forgetting to equations of the first degree as to more complicated logarithmic equations. These curves are more or less parallel, more or less approaching linear regression, and as often tending to positive as to negative acceleration. One thing we can definitely state is that, on the whole, they are not logarithmic.

It was pointed out in the last chapter that relearning based on the "saving method" produces a type of curve vastly different from the results of the other methods which measure the amount of reproduced material. Now it becomes further evident that among the latter group, recognition *sometimes* has its unique curves which are as different from those of the other methods as relearning is from all the rest. Under certain circumstances it may happen that each of these curves will take on a logarithmic form. Strong's curve for recognition memory, for instance, is like Ebbinghaus' curve for the "saving method," and the former can certainly be used as illustrative of Bean's generalized statement. We do not maintain that relearning, anticipation, written reproduction, recognition, reconstruction, etc., each has a general curve. We only indicate that the logarithmic assumption and even the phenomenon of negative acceleration may totally disappear upon further investigation.

However, if the problem be put in such a way that we have to choose between Ebbinghaus and Ballard, then the former type of curve is much closer to our own results.

IV. THE EFFECT OF EXTENDING THE TIME LIMIT FOR RECALL UPON THE AMOUNT OF MATERIAL RECALLED

In written reproduction, as described above, a record was taken at the end of the first minute of recall and another at the end of the second minute, when the subject did not finish recalling within these time limits. These preliminary records, together with the other records which were completed within 1 or 2 min., may be taken as indicative of the proficiency of written reproduction memory up to 1 and 2 min. of recall respectively.

In recognition similar preliminary records were taken at the end of 90 sec. without the subject's knowledge. The difference between these 90 sec. records and the complete records has been referred to in Ch. II.

We present in Table XVIII the comparative results for the written reproduction methods for four degrees of the original learning. The values for 100% learning are averages from both series of experiments. When the same written reproduction records were scored for position and sequence, the preliminary values were slightly changed. These values are tabulated in Table XIX to facilitate comparative study. Similar results for the recognition method are presented in Table XX.

1. Comparison of the Written Reproduction Results.

It was indicated in connection with a discussion on anticipation and relearning that partial retention is not as readily recalled as complete retention. Within certain limits, the duration of the recalling process may directly correspond to the strength of the association. However, the tendency for the amount of recall to increase upon extending the time limit very soon becomes ineffective. One conclusion we can draw from Tables XVIII and XIX is that the effect of extending the time limit for recall with the written reproduction method becomes less and less important beyond 1 or 2 min. The extension from 1 to 2 min. increases the scores by a far greater amount than does a further extension from 2 to 5 min. This particular effect is quite independent of minute variations in the method of measurement. Scoring position and sequence does not in the least change the relative importance of the successive extensions of the time limit. The

TABLE XVIII

WRITTEN REPRODUCTION, COMPARISON OF PRELIMINARY AND FINAL RECORDS

	20 min.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 day	2 days
150% learning									
1 min.			75.9	60.0	57.9	43.0	34.6	26.0	22.9
Difference..			8.2	16.5	13.1	13.7	8.7	5.7	3.1
2 min.			84.1	76.5	71.0	56.7	43.3	31.7	26.0
Difference..			3.9	7.9	10.9	8.9	11.1	6.8	4.8
Complete ...			88.0	84.4	81.9	65.6	54.4	38.5	30.8
Total D....			12.1	24.4	24.0	22.6	19.8	12.5	7.9
100% learning									
1 min.	75.6	69.2			44.3			30.8	26.4
Difference..	10.5	10.8			11.5			7.3	3.8
2 min.	86.1	80.0			55.8			38.1	30.2
Difference..	3.3	4.0			6.8			4.3	3.3
Complete	89.4	84.0			62.6			42.2	33.5
Total D....	13.8	14.8			18.3			11.6	7.1
67% learning									
1 min.	69.8	47.7			42.5			31.7	21.5
Difference..	12.7	14.0			16.2			6.6	2.7
2 min.	82.5	61.7			58.7			38.3	24.2
Difference..	2.9	10.8			7.1			3.1	.6
Complete	85.4	72.5			65.8			41.4	24.8
Total D....	15.6	24.8			23.3			9.7	3.3
33% learning									
1 min.	59.8	43.1			31.2			21.5	12.7
Difference..	6.2	7.1			8.8			4.5	.4
2 min.	66.0	50.2			40.0			26.0	13.1
Difference..	1.7	3.8			2.7			.2	.6
Complete	67.7	54.0			42.7			26.2	13.7
Total D....	7.9	10.9			11.5			4.7	1.0

latter also holds for all degrees of the original learning except for the 12-hr., 1-day and 2-day intervals with 150% learning. In these exceptional cases the second extension of three minutes seems to be more effective than the first extension of one minute. Two of these intervals, 1 day and 2 days, correspond to the points

TABLE XIX

WRITTEN REPRODUCTION, COMPARISON OF PRELIMINARY AND FINAL RECORDS, SCORING POSITION AND SEQUENCE

	20 min.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 day	2 days
150% learning									
1 min.			72.4	56.7	53.3	38.2	32.4	22.9	20.0
Difference..			8.5	16.4	11.7	12.5	8.2	4.3	1.8
2 min.			80.9	73.1	65.0	50.7	40.6	27.2	21.8
Difference..			3.8	9.3	10.0	7.9	10.1	5.0	2.4
Complete			84.7	82.4	75.0	58.6	50.7	32.2	24.2
Total D....			12.3	25.7	21.7	20.4	18.3	9.3	4.2
100% learning									
1 min.	71.1	64.4			39.4			26.3	22.9
Difference..	9.4	8.9			9.0			5.7	3.0
2 min.	80.5	73.3			48.4			32.0	25.9
Difference..	2.9	3.9			5.8			3.0	2.2
Complete	83.4	77.2			54.2			35.0	28.1
Total D....	12.3	12.8			14.8			8.7	5.2
67% learning									
1 min.	65.1	42.0			36.4			26.1	17.2
Difference..	12.8	12.5			13.6			5.2	2.5
2 min.	77.9	54.5			50.0			31.3	19.7
Difference..	2.6	9.6			7.4			2.2	.3
Complete	80.5	64.1			57.4			33.5	20.0
Total D...	15.4	22.1			21.0			7.4	2.8
33% learning									
1 min.	53.3	37.7			26.2			17.7	9.5
Difference..	6.1	5.7			6.8			3.6	.2
2 min.	59.4	43.4			33.0			21.3	9.7
Difference..	1 3	4.2			2.2			.2	.8
Complete	60.7	47.6			35.2			21.5	10.5
Total D....	7.4	9.7			9.0			3.8	1.0

where the retention values for the 150% learning curve decrease very rapidly, as noticed in the last chapter. The facts still await an explanation, but we shall have to postpone further discussion to the end of Ch. V.

With increase in the length of the time interval, the same decline in the effect of extending the time limit is also observed. Previous to the lapse of the 4-hr. interval, the increment due to extending the time limit grows larger and larger. It then suddenly decreases with time. This characteristic change of the increment prevails under various conditions. It holds for all degrees of learning. It is not altered by scoring position and se-

TABLE XX

RECOGNITION, COMPARISON OF PRELIMINARY AND FINAL RECORDS

	20 min.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 day	2 days
150% learning									
90 sec.			96.7	94.1	91.1	91.5	88.7	85.8	72.8
Complete			97.5	95.8	93.3	91.6	92.5	83.2	72.8
Difference..			.8	1.7	2.2	.1	3.8	—2.6	.0
100% learning									
90 sec.	95.7	93.8			91.5			71.2	69.4
Complete	96.8	94.8			92.7			76.1	75.2
Difference..	1.1	1.0			1.2			4.9	5.8
67% learning									
90 sec.	92.9	94.6			85.2			72.6	61.5
Complete	93.2	93.3			84.7			73.7	61.5
Difference..	.3	—1.3			—.5			1.1	.0
33% learning									
90 sec.	75.0	66.0			54.3			44.6	29.4
Complete	73.3	64.4			54.6			45.7	25.5
Difference..	—1.7	—1.6			.3			1.1	—3.9

quence, nor is it minimized when the increment is computed as a percentage of the total score.

The strength of retention for the different syllables is, therefore, not equal. Some are more easily recalled than others, and consequently take less time. When the time interval is lengthened, the more difficult and uncertain ones deteriorate first. The amount of this partial retention increases in value for the first few hours after learning. Thereafter, not only does retention as a whole deteriorate with time, but the strength and the amount of partial retention also decrease so that its re-instatement upon the extension of the time limit becomes less and less probable. The assumption as presented in Ch. II which states that the amount of partial retention increases with the time interval for all memory processes is evidently unsound.

We can mention the fact in this connection that from 4 hours to 1 day is a long interval, and that in the latter case, sleep occurs between learning and relearning. Increase in the degree of the original learning seems to arrest this change in the magnitude of the increment, though very ineffectively. Thus, with 150% learning there seems to be a plateau in the effect of extending the time limit for the intervals from 3 to 12 hours inclusive, and then the sudden decrease. This would mean that the growth of the total amount of forgetting is arrested to that extent. In 100%, 67% and 33% learning, the abruptness of that change which occurs at the lapse of the 4-hr. interval to the increment resulting from longer durations of the recalling process is also somewhat proportional to the degree of the original learning.

2. *Comparison of the Recognition Results.*

Coming now to recognition, the effect of extending the time limit beyond 90 sec. is quite different from the results we just discussed. In some respects, the conclusions from these two methods of measurement are contradictory. If we can directly compare the magnitude of the scores, the increment for recognition is very much smaller. But a more significant contrast is that, with 100% learning, the increment for recognition is almost constant in value until the lapse of the 4-hr. interval, and then suddenly increases for the 1-day interval and becomes still higher at the end of 2 days. This contradicts what we have found in written reproduction in every particular. Assuming that partial retention takes more time for recall, as we have done, these characteristics of the recognition process seem to corroborate Strong's[2] conclusion which maintains that in recognition memory the amount of partial recognitions does not decrease in time as fast as the amount of recognition as a whole. One may even conclude that the partial recognitions actually increase with the length of the interval.

With 150% learning there seems to be a general increase of the effect of extending the time limit upon the amount of recognition until the lapse of the 12-hr. interval. At that point the increment disappears. The subjects seemed to have selected all the syllables they could recognize at the end of 90 sec., the rest of the process being chance performance. It is difficult to think of an increase of 50% in the original learning as affecting a decrease in the amount of recognition. A tentative explanation of this fact will be presented at the end of Ch. V.

For 67% and 33% learning, there is no sudden change of

[2] *Op. Cit.*, pp. 352 ff.

the increment as with 150% Extending the time limit for recognition does not result in any appreciable change in the numerical values either way, except that at the end of the 2-day interval for 33% learning, it affects a considerable decrease in the total score. This is to be expected if we remember that finally the amount of partial recognition may itself decrease so that even the presentation of the original material does not avail. Nevertheless, the subject was required to complete the selection of 12 syllables, depending upon chance; hence, there was a decrease in the total score. This final decrease in the amount of partial recognition is also hastened by a less complete degree of learning. Thus, while the numerical value for the 2-day interval with 67% learning is not changed by the extension of the time limit for recall, the corresponding score was decreased by 3.9 with 33% learning.

3. The Gradation of the Memory Processes.

The results from both methods of measurement seem to indicate that written reproduction and recognition memory are quite different, but the difference is still one of degree, not of kind. One process may pass over to the other. Judging from commonsense, recognition memory lasts longer than the ability to reproduce *ad verbatim*. What fa ls even the vaguest recall may, upon the presentation of the object, flash into distinct recognition. On the other hand, if enough time be allowed, one can as a rule recognize what he can recall. It seems that the memory processes are graded in some such way as the following:

A. Complete retention.

B. Partial retention, which takes time for recall, and may involve errors, as will be seen in another chapter.

C. Still less permanent retention which may completely escape written reproduction, but which, nevertheless, may be reinstated in recognition.

D. Partial recognition memory which can be developed upon extending the time limit for recall.

E. Retention that cannot be measured even by the method of recognition. It approaches complete obliviscence and occurs at the end of a fairly long interval after incomplete learning.

Condition A gives place to B soon after learning, at least with a great part of the retained material. Thereafter B approaches C faster than A does B so that the total amount of B decreases with the lapse of the 4-hr. interval. C lasts for a long time and gives place to D. D is not effective upon the immediate presentation of the original material, but may return in the process of recognition. When the degree of the original learning is only 33%, the amount of D becomes insignificant at the end of 2 days. The paradoxical effect is to shorten the duration of the actual process of recognition to less than 90 sec. when the syllables that can be recognized have all been selected, but apparently to increase the duration beyond that time limit. Nothing being remembered, the extended time serves only to fulfill the requirement of the experiment, which is to select 12 syllables.

4. The Effect of Extending the Time Limit upon the Shape of the Curves of Retention.

(1) In written reproduction the curve changes with the magnitude of the increment resulting from extending the time limit to 5 min. As compared with the preliminary curve for the 1- or 2-min. records, the final curve falls more gradually within the first four hours after learning, and then more suddenly. If the preliminary curve of forgetting be logarithmic, the effect of the increment would at least tend to complicate the function.

(2) In recognition with 67% learning, the more gradual fall of the curve owing to the extension of the time limit continues until the end of the second day.

(3) In recognition with 67% and 33% learning, extending the time limit does not bring about any appreciable change in the curves.

(4) *The shapes of the curves are practically determined for written reproduction at the end of 2 min. and for recognition at the end of 90 sec.*

V. THE RELATION BETWEEN THE AMOUNT OF ERROR AND OTHER FACTORS

In the recognition and reconstruction methods of measurement, the score was determined directly by the amount of error made in recall as well as by the amount of correct material. With the "saving method" the number of presentations required for relearning was determined by the amount of forgetting, but incidentally also by the amount of error. But errors were disregarded in the scores for anticipation and written reproduction. An attempt to keep separate records for the number of incorrect responses in anticipation met with failure. The duration of recall for each syllable was too short, considering the fact that the experimenter already had to turn the drum and to record success or failure between responses. In written reproduction, however, the records were permanent and we could study the amount and nature of error after the experiments were completed.

1. The Amount of Error as a Function of the Length of the Interval.

The amount of error made in written reproduction for each time interval and each degree of learning is presented in Table XXI. With the possible exception of 33% learning, the amount of error increases with the length of the interval.

TABLE XXI

AMOUNT OF ERROR IN WRITTEN REPRODUCTION

	20 min.	1 hr.	2 hrs.	3 hrs.	4 hrs.	6 hrs.	12 hrs.	1 d.	2 d.
150% learning...			2.7	3.1	6.7	7.3	8.3	12.9	15.8
100%	4.2	5.8			11.7			11.7	11.3
67%	3.3	4.6			5.4			7.7	9.0
33%	4.4	7.9			10.8			5.2	8.7

2. The Relation between the Amount of Error and the Degree of Learning.

In Table XXI the amount of error for 150% learning is probably smaller than for any other degree of learning with the shorter intervals, but it grows to be the largest when the interval is lengthened.

On the whole, error does not seem to be proportional to the degree of learning until the lapse of the 4-hr. interval. Thereafter the gross amount of error increases with the degree of learning.

These facts may be compared with the effect of extending the time limit upon the amount of reproduction, as discussed in Ch. IV. The magnitude of that effect for 150% learning changes from the shorter to the longer intervals in the same manner as does the amount of error. Before the lapse of the 4-hr. interval also, that effect is not proportional to the degree of learning, but the two factors take on a functional relationship for the 1-day and 2-day intervals. So the relationship between learning and error is similar to the relationship between learning and the effect of extending the time limit. However, the latter effect decreases with the lapse of the 4-hr. interval, while the amount of error increases with the time interval to the end of 2 days.

As already explained, the effect of extending the time limit is a function of the amount of partial retention. It is proportional to the amount of retention of condition B (p. 60) which can be reinstated in written reproduction but which takes time. Now we are ready to state a theory as to the significance of error making, and to see how it can be applied to explain the facts enumerated. The condition of error making is not complete forgetting or obliviscence. *It is the presence of partial retention* which can hardly be reinstated but which, nevertheless, is so near the point of complete recall as to cause conflict and confusion. Errors are made mostly in the change from condition B to C.

Thus, with 150% learning, on account of the higher degree of original mastery, the amount of partial retention of condition B is small as compared with that of A. The amount that is due to the change from condition B to C is also insignificant. Hence, we have the smaller amount of error for the short intervals. When the time interval is lengthened, condition B prevails and the amount of error increases steadily.

Similarly, we can explain why the amount of error is parallel to the effect of extending the time limit in almost every instance. The two are reciprocal functions and both are due to the existence of partial retention. Eventually the amount of error will decrease with the amount of partial retention, but the former may keep on increasing while the effect of extending the time limit has reached a climax. The amount of error increases because condition B approaches C much nearer for the longer intervals than for the shorter ones. Therefore, the increase in score due to extending the time limit, plus the amount of error, is an approximation of the amount of partial retention. The change from condition B to C is a very complicated process.

3. The Relation between the Amount of Error and the Amount of Retention.

When the time interval is lengthened, retention decreases but error increases. So if the amount of error is calculated as a percentage of the amount of retention, the values will increase more rapidly. The significance of this comparison is questionable. The amount of error should rather be compared with the amount of forgetting, the whole syllable series minus the amount of retention, which therefore increases with the time interval.

4. The Relation between the Amount of Error and the Amount of Forgetting.

The comparative data for the amount of forgetting and of error are presented in Table XXII.

The increase with the time interval in the amount of error is much slower relative to the increase in the amount of forgetting. This difference in the rate of increase with the time interval is further inversely proportional to the degree of learning.

As already stated, errors are made mostly in the transition from condition B to C, and the partial retention that cannot be even thus reinstated is apparently forgotten, according to the written reproduction method of measurement. So with the increase in the time interval, the amount of forgetting has a higher accumulative value than is possible to the amount of error. The former naturally increases more rapidly. Further, if a higher degree of

TABLE XXII

COMPARISON OF THE AMOUNT OF FORGETTING AND OF ERROR IN WRITTEN REPRODUCTION

	20 m.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 d.	2 d.
150% learning									
Forgetting			12.0	15.6	18.1	34.4	45.6	61.5	69.2
Error			2.7	3.1	6.7	7.3	8.3	12.8	15.8
100% learning									
Forgetting	9.4	14.2			35.2			54.4	59.8
Error	4.2	5.8			11.7			11.7	11.3
67% learning									
Forgetting	14.6	27.5			34.2			58.5	75.2
Error	3.3	4.6			5.4			7.7	9.0
33% learning									
Forgetting	32.3	46.0			57.3			73.8	86.3
Error	4.4	7.9			10.8			5.2	8.7

learning tends to arrest the whole process of forgetting, then the difference between the increase in the amount of forgetting and of error will be inversely proportional to the degree of learning.

Aside from these general statements, the relationship between error and forgetting is obscure. With 150% learning, the amount of forgetting is to the amount or error as 5 is to 1, for all the time intervals. This ratio regularly decreases with the time interval when the degree of learning is reduced to 67%. With the other degrees of learning, there is no definite relationship. The P.E.'s of the values that constitute the curve for the amount of error are too high to give warrant to further generalization.

However, this lack of similarity or causal relationship between the two groups of values only intensifies the problem. The amount of error is indicative of the amount of partial retention. In other words, the curve for the amount of error is a part and parcel of the curve of forgetting. Upon the evaluation of this relationship will depend whether a universal mathematical statement of the problem is possible.

THE EFFECTS OF 150% LEARNING UPON VARIOUS FACTORS

From the above section, it is found that the relationship between forgetting and error is more definite for 150% than for any other degree of learning. The ratio of the two amounts does not decrease with the time interval as it does with lower degrees of learning. We are now ready to gather together what has been noticed in the last two chapters concerning the characteristics of 150% learning.

1) Its retention values are lower than for 100% learning at the end of 1 and 2 days in written reproduction, at the end of 1 day in reconstruction and at the end of 2 days in recognition (p. 50).

2) The effect of extending the time limit upon the amount of written reproduction is peculiar in this case. A second extension of the time limit from 2 to 5 min. brings about a larger increment than the first extension from 1 to 2 min. This relative effectiveness of the two extensions of the time limit is just the reverse of what has been found with all other degrees of learning (p. 56).

3) Extending the time limit for recognition gives negative results at the end of the 1- and 2-day intervals. The effect of the extension is different for the shorter intervals with the same degree of learning and for all the intervals with 100% learning (p. 59).

An explanation for these peculiarities is possible under the following suppositions.

A) A high degree of learning, as we have indicated, tends to arrest the process of forgetting. In the process of forgetting, retention of condition A approaches partial retention of condition B, the latter approaches C, etc.

B) Most of the syllable series for 100% learning were given each subject before the 150% series. In the former experiments they had noticed that the amount of retention after 1 or 2 days was comparatively low. This fact might have influenced their attitude when they came to the longer intervals in 150% learning.

Combining these two postulates, it seems probable that for the longer intervals there may be a conflict of condition B with both A and C, when B is intensified with over-learning. The retroactive effect is (1) to decrease the total amount of written reproduction. (2) It naturally follows that the effect of the second extension of the time limit will be greater than that of the first, for the conflict will tend to prolong the recall process of even the well retained members. (3) And as condition B always approaches C and also conflicts with C, the amount of error will be increased for the longer intervals.

Even on this hypothetical basis, there is no explanation for the sudden drop of the retention curve for recognition at the end of 2 days. In the same connection the extension of the time limit gives negative results. Probably, over-learning causes a part of the syllabie series to stand out more distinctly, thus contrasting with the relative "amnesia" of the other parts.

Were these explanations unsatisfactory, the fact remains conclusive that over-learning to the extent of 150% is at less advantage than 100% learning when tested for retention at the end of 1 and 2 days.

VI. THE DURATION AND THE SPEED OF THE PROCESS OF RECALL IN WRITTEN REPRODUCTION, RECOGNITION AND RECONSTRUCTION

Two scores, one of amount and one of time, are not mutually interpretative, but the latter may in a way be indicative of the nature and the strength of retention, though not in quantitative terms. The data thus far presented all concern the amount of retention, correctly or incorrectly reproduced. In Ch. IV we discussed the effect of different time limits only in relation to the amount of reproduced material. We can now consider the duration and the speed of the process of recall with the different methods of measurement.

The "saving method" does not give separate measurements for time and amount, so the two cannot be independently treated. The duration or the speed of the anticipation process is not significant, its value being a constant, 2 sec. for each syllable.

With the other three methods of measurement, duration and speed are quite independent of the amount of retention. In written reproduction the subject could take any length of time up to 5 min. As a rule, he did not take as long as 5 min. to finish the process. In recognition and reconstruction, there was not even a time limit. So the relationship between the time and the amount of recall becomes an important problem.

1. Duration as a Function of the Order in which the Measurements were Taken.

In Table XXIII are presented the average durations of the three processes of recall in number of seconds.

From Table XXIII, written reproduction had the longest duration, recognition the second and reconstruction the last. That was exactly the order in which the measurements were taken, written reproduction being the first, recognition the second and reconstruction the last.

TABLE XXIII

DURATION OF RECALL, NO. OF SEC.

	20 m.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 d.	2 d.
Writ. Reprod.									
150% learning			134.4	171.4	193.8	209.5	198.6	186.4	178.7
*100% A......	112.2	134.0			147.9			132.1	97.0
100% B......	145.5	171.2			211.6			182.7	176.0
67%	139.0	207.4			189.3			157.5	154.6
33%	108.8	155.1			124.2			108.3	101.4
Recognition									
150% learning			60.6	71.7	81.8	94.3	99.9	86.9	114.2
*100% A......	76.7	87.1			108.1			126.9	132.8
100% B......	72.6	78.4			92.6			89.9	101.3
67%	68.9	83.2			98.9			92.7	107.9
33%	88.4	104.0			100.6			93.1	114.0
Reconstruction									
150% learning			29.5	32.3	41.0	50.2	51.1	61.4	60.2
*100% A......	47.9	41.9			50.4			62.3	68.7
100% B......	50.0	41.1			45.3			57.9	59.2
67%	37.5	53.2			48.2			58.0	57.8
33%	43.7	44.5			52.3			48.1	51.7

*A The first series of experiments.
B The second series of experiments.

The duration of written reproduction was shorter in the first series of experiments than in the second series, but the comparative duration of recognition was *vice versa*. No important difference between the two series of experiments was recorded in reconstruction.

So the difference in duration may have incidentally resulted from the technique of the experiments, quite independent of the relative difficulty of the processes. Each recall process may have facilitated the succeeding one. The longer the written reproduction process, for instance, the shorter was the duration of recognition. However, as we have seen in Ch. II, the conditions of recall differed in the number of restricting factors involved. Most probably, these factors also influenced the duration of the processes.

2. The Speed of Recall.

The recognition and reconstruction methods of measurement required the subject to complete the process of recall regardless of the amount of correct retention. So with these two methods, the duration and the speed of recall are identical. The conditions were very different in written reproduction. There the subject could "give up" at any time. The averages presented in Table XXIII fall far below the limit of 5 min. The average subject would "give up" long before the lapse of that time limit. At the same time, the amount of reproduction varied with the amount of retention, and the differences in duration might simply be in part a function of the amount of material that was reproduced.

To avoid this difficulty, the speed of written reproduction is calculated as the number of seconds per unit material. If the duration of recall varied only as an effect of the amount of reproduction, then the speed per unit would have a constant value for all the intervals. Such, however, is not the case, as may be seen in Table XXIV.

TABLE XXIV

SPEED OF REPRODUCTION, NO. OF SEC. PER UNIT MATERIAL

	20 m.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 d.	2 d.
150% learning			1.9	2.4	2.6	4.2	4.8	6.9	8.6
100%	1.7	1.1			3.8			6.0	7.3
67%	1.8	3.5			4.4			6.3	9.3
33%	1.8	3.2			3.6			6.1	11.1

Comparing that table with the recognition and reconstruction values in Table XXIII (except the data from the 1st series of experiments), the relationship between the speed of recall and other factors may be summarized as follows:

(*1*) *The speed of recall decreases with the time interval.* This generalization holds for written reproduction without exception. That such a tendency exists in recognition is also unquestionable. With reconstruction, the difference in speed between the 1-day and 2-day intervals is very slight, and there are also marked deviations from the assumed functional relationship for the shorter intervals. But the decrease in speed with time may be clearly seen if we taken the average of all the degrees of learning.

(*2*) *On the whole, the speed of recall increases with the degree of learning.*

(*3*) Since the duration of recognition and reconstruction increases with the time interval, while the amount retained and the accuracy of the two processes decrease with time, it follows that *speed increases with accuracy*. The increase in speed with respect to accuracy is much faster than the increase with respect to the time interval.

3. Comparison between the Speed of Recall and the Amount of Forgetting.

The increase with the time interval in the number of seconds per unit recall, as further complicated by the degree of learning, may be compared with the amount of forgetting which also increases with time and varies with the degree of learning. The data are gathered together in Table XXV. It is useless to compare the absolute amount of forgetting with the total speed of the process, since only in the light of the general shape of the curves can such a comparison be intelligible. The values presented in Table XXV are converted from the original data, using the numerical value obtained from the shortest interval in each case as the unit. The table reveals especially the relative increase in the values with the length of the time interval.

In spite of the arbitrary process in reducing the amount of forgetting and the speed of recall to the same unitary basis, any graphical representation of the data in Table XXV would still be so complicated as to make interpretation impossible. The curves of forgetting may be said to be logarithmic in a sense, but a similar mathematical statement would no longer hold for the curves for the speed of recall. A possible generalization one can make from these facts is that the curves of speed do not rise as rapidly as the curves of forgetting. The decrease in the speed of recall with respect to the length of the time interval is not as rapid as the decreases in the total amount of retention.

So the amount of forgetting and the speed of recall do not have the same type of curves. The results of this section agree with the conclusion reached in Ch. II concerning the difference between the relearning and the other methods, only we have attained the additional observation that the time curves differ among themselves even more radically than do the curves of for-

TABLE XXV

COMPARISON OF THE AMOUNT OF FORGETTING WITH THE SPEED OF RECALL, TAKING THE NUMERICAL VALUE OF THE SHORTEST INTERVAL AS UNIT

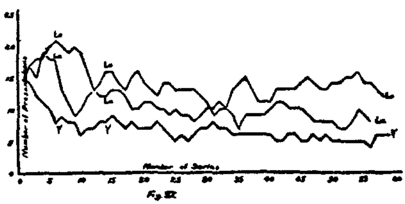
	20 m.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 d.	2 d.
Writ. Reprod.									
150% learning									
Forgetting..			1.00	1.30	1.51	2.87	3.80	5.12	5.77
Speed			1.00	1.26	1.37	2.21	2.53	3.63	4.53
100% learning									
Forgetting..	1.00	1.51			3.74			5.78	6.36
Speed	1.00	1.24			2.24			3.53	4.29
67% learning									
Forgetting..	1.00	1.88			2.34			4.01	5.15
Speed	1.00	1.94			2.44			3.50	5.17
33% learning									
Forgetting..	1.00	1.42			1.77			2.28	2.67
Speed	1.00	1.78			2.00			3.39	6.17
Recognition									
150% learning									
Forgetting..			1.00	.63	.74	1.70	1.50	4.53	4.49
Speed			1.00	1.18	1.35	1.56	1.65	1.43	1.88
100% learning									
Forgetting..	1.00	.90			2.35			4.80	5.23
Speed	1.00	1.08			1.28			1.24	1.40
67% learning									
Forgetting..	1.00	2.76			4.59			5.42	8.52
Speed	1.00	1.21			1.44			1.35	1.57
33% learning									
Forgetting..	1.00	1.56			2.13			3.03	3.28
Speed	1.00	1.18			1.14			1.05	1.29
Reconstruction									
150% learning									
Forgetting..			1.00	1.68	2.68	3.28	3.00	6.72	10.88
Speed			1.00	1.37	1.39	1.70	1.73	2.08	2.04
100% learning									
Forgetting..	1.00	1.19			2.00			5.33	5.02
Speed	1.00	.82			.91			1.16	1.18
67% learning									
Forgetting..	1.00	.99			2.25			3.87	5.66
Speed	1.00	1.42			1.29			1.57	1.57
33% learning									
Forgetting—	1.00	1.33			1.70			2.03	2.79
Speed	1.00	1.02			1.30			1.10	1.18

getting. If the relearning method could be varied and controlled as are the other methods, the amount of "saving" would most probably increase or decrease according to various conditions

VII. INDIVIDUAL DIFFERENCES AND CORRELATIONS

1. *Individual Differences in Practice Effect.*

It has been sometimes maintained that the practice effect in learning nonsense material disappears after the successive mastery of but a few series. In our experiments we had the opportunity to plot the practice curve for each individual subject. About one-half of the subjects learned more than 50 series each, thus giving rather extensive practice curves. From the comparison of these curves, it seems that the extent of practice effect is subject to individual differences. For most of the subjects it is present even after the mastery of 40 or 50 series, and it decreases with different rates for different individuals. Three typical curves are given in Fig. XI (Smoothed).



Generally speaking, the practice effect in these curves seems to have disappeared after the tenth trial, but individual differences are, nevertheless, present. The curve for subject Y has an abrupt initial drop which is not so prominent in the other curves. That is, he adapted himself to the situation more quickly than did the other subjects. The readiness with which one adapts to this particular situation is in a way proportional to his speed of learning. After the tenth trial the gradual fall of the curves for Le and Y is noticeable to the very end, but the curve for Lo is stationary. So individuals differ not only in the amount of the initial drop, but also in the gradual decline of the practice effect. A break occurred in the curve for Y when for more than two months the subject did not work on nonsense material. The practice effect was carried over that interval.

In Ch. III it was mentioned that the subjects could use 67% of the average number of presentations to a greater advantage after they had practiced in 150% and 100% learning for more than 5 weeks. This improvement in learning also varied greatly according to the individual. As previously stated, we took an average of the number of presentations required by each individual for the mastery of the 20 syllable series in 100% and 150% learning. The number of series that we learned with only two-thirds or less of the average number of presentations, as determined for each individual, was exceedingly small. However, when they came to 67% learning, most of the subjects learned one or more of the series with that number of presentations. Table XXVI shows how the individuals differ in this respect.

TABLE XXVI

SERIES LEARNED WITH 2/3 OR LESS OF THE AVERAGE NUMBER OF PRESENTATIONS

SUBJECT	In the first 20-25 series	In the next 10
B	1	3
I	0	1
Ka	2	3
Ko	0	1
Lud	2	1
Luh	1	2
S	2	4
Wi	0	1
Wo	1	2
Y	0	0

2. Individual Differences in the Speed of Learning.

Table XXVII presents the average number of presentations required by each subject to learn a series of 12 nonsense syllables.

It takes an average subject about 14 presentations to learn a series of 12 nonsense syllables by the anticipation method. This average is just a little higher than that given by Finkenbinder.[1]

The two groups are about equal in efficiency, for as already

[1] *Op. cit.*, pp. 21-22.

TABLE XXVII

AVERAGE NUMBER OF PRESENTATIONS FOR EACH SERIES

SUBJECT	NO. OF PRESENTATIONS	P. E.	NO. OF SERIES LEARNED
*Luh	4.95	.20	20
*Y	6.42	.18	40
Ko	6.85	.26	20
*I	8.75	.33	20
*Wa	9.42	.27	50
*Le	9.84	.21	50
*C	10.90	.55	20
R	11.88	.32	50
DW	13.26	.37	50
Lo	13.50	.28	50
Lud	13.65	.33	20
Ka	16.20	.57	20
*S	16.50	.62	20
Wi	17.20	.37	20
*Ts	18.10	.54	20
B	18.60	.51	20
Kan	19.75	.86	20
Wo	19.80	.65	20
*F	24.50	.67	20
Av.	13.69		

*Chinese students.

indicated, it is more difficult to spell a syllable letter by letter as required in our experiments than to pronounce it as a whole as required by Finkenbinder. A group of 15 to 20 subjects is, therefore, large enough to make a random sample.

The fastest learner of the group is five times as proficient as the slowest learner. On the whole, the Chinese students are better memorizers than the Americans. The averages stand as 12.15 against 15.07. With the exception of subject F, the difference would be much higher. Six Chinese occupy the first seven places. This superiority of the memorizing ability of the Chinese is interesting in connection with the problems of classical training and the improvement of memory with practice.

3. Individual Variability in the Amount of Retention.

Applying the formula $V = \frac{100\ S.\ D.}{\text{Mean}}$, we calculated the coefficient of variability of the average amount of retention for the five methods of measurement and the four degrees of learning. These coefficients are tabulated in Table XXVIII.

XXVIII

Coefficients of Variability

	20 m.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 d.	2 d.
Anticipation....	26.2	28.9			41.0			80.9	133.3
Relearning	18.1	31.8			22.6			31.3	34.3
Writ. Reprod.									
150% learning			15.2	17.5	12.5	26.9	33.1	77.1	76.3
100%	6.8	11.7			27.0			56.8	70.4
67%	6.6	28.5			31.2			48.1	106.7
33%	17.8	32.0			38.5			86.4	115.5
Recognition									
150% learning			3.4	4.4	6.6	9.0	15.2	18.5	18.9
100%	4.3	5.8			7.8			22.9	20.9
67%	11.3	8.6			13.2			25.6	31.8
33%	19.3	16.6			41.6			62.9	66.5
Reconstruction									
150% learning			15.6	12.2	9.8	14.8	23.8	53.9	69.8
100%	10.3	12.7			17.0			46.3	57.0
67%	8.7	21.3			23.8			28.7	70.3
33%	22.4	32.3			41.4			85.5	78.4

The magnitude of the coefficients of variability for the long intervals indicate that the data are not statistically reliable. However, our present problem is exactly to describe, if not to interpret, the differences in statistical reliability, or individual variability, under certain conditions. So the above statement directly leads to the following generalization.

(1) *Individual variability increases with the time interval.* This factual statement holds for all the methods of measurement and all degrees of learning. The tendency can be more distinctly observed as graphically represented in Fig. XII, for 100% learning.

These graphs are based on similar, and in some cases the same numerical values as are the retention curves in Fig. II. A comparison of the figures brings out the fact that the curves for the coefficients of variability have not only preserved the particular order of the different methods of measurement, a proportionality of the absolute differences in numerical value, but to some extent, even the specific similarity or dissimilarity in general shape;

they also manifest that characteristic difference between relearning and the other methods of measurement. The relearning curve begins as the second in height in the order of variability but ends as the fourth, having crossed the written reproduction and reconstruction curves and now stands half way between recognition and reconstruction both in shape and numerical value. The significance of this comparison has been discussed at length in Ch. II, where it was pointed out that relearning is not a helpful method for the study of individual differences.

In the present connection, the coefficients of variability for the different intervals as determined by the relearning method fall within a close range in numerical value. Just as the retention curve for relearning is complicated by involving two learning curves for memory, so also the corresponding curve for the coefficients of variability becomes almost unanalyzable because of the same complicaion. In Fig. XII the asterisk represents the

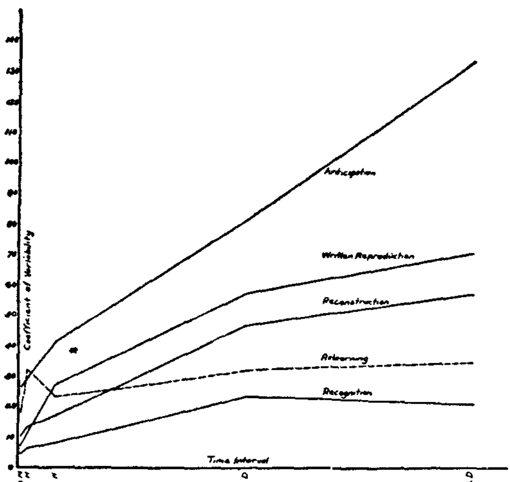


Fig XII

coefficient of variability for the original learning, which is 37.8. The implications of the "saving method" would lead one to expect that this value would be rather close to the average value of the relearning curve in the same figure.

(2) *Individual variability increases inversely as the degree of*

learning. Exceptions are found in written reproduction and reconstruction with 150% learning at the end of the 1-day and 2-day intervals. The explanation given in Ch. V for the characteristics of 150% learning will cover the present instances.

That the increase in the degree of learning should tend to equalize and finally to eliminate individual differences is also to be expected. For a certain degree of over-learning will make the material so well fixed for any individual who can fulfill the requirements of these experiments that the amount of retention will be always 100%. Such are our habitual and conventionalized reactions as we so often observe in daily life. The variability of these reactions approximates 0. In other words, the effect of increasing the degree of learning upon the coefficient of variability approaches 0 as a limit.

As a corollary to the above conclusion, the phenomenon of diminished returns may be described more specifically. Since the increase in the degree of learning brings about diminished returns in the amount of retention and at the same time neutralizes individual variability, it follows that, in the long run, an individual very efficient in immediate retention is not likely to improve with over-learning as rapidly as another who is very inefficient in immediate retention.

(3) *Individual variability also seems to be a function of the method of measurement and to increase with the number of restricting factors involved in the recall process,* as explained in Ch. II.

(4) Summarizing (1), (2) and (3), *individual variability in retention increases with the difficulty of the act and decreases with the frequency and the recency of practice.*

4. Correlation between the Speed of Learning and the Amount of Retention.

The formula used is

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)}. \quad \text{Then } r = 2 \sin \left(\frac{\pi}{6}\rho\right)$$

and P. E. $= .703 \dfrac{1 - r^2}{\sqrt{N}}$. Table XXIX gives the correlations whose numerical values are at least twice as large as their respective P. E.'s.

TABLE XXIX

Correlation between the Speed of Learning and the Amount of Retention

	20 m.	1 h.	2 h.	3 h	4 h.	6 h.	12 h.	1 d.	2 d.
Relearning		—.42						—.42	
P. E.		.18						.18	
Anticipation									.51
P. E.									.18
Writ. Reprod.									
150% learning				.43					
P. E.				.18					
100%	.30	.40							
P. E.	.15	.14							
67%	.40								
P. E.	19								
33%								.50	
P. E.								.17	
Recognition									
150% learning			.59	.53		—.47			—.64
P. E.			.14	.16		.17			.13
100% learning	.39	.32			.37				
P. E.	.14	.15			.14				
67% learning	.42							—.40	—.57
P. E.	.18							.19	.15
33%									
P. E.									
Reconstruction									
150% learning			.50				.44		
P. E.			.17				.18		
100% learning									
P. E.									
67% learning	.58								—.47
P. E.	.15								.17
33% learning	.41								
P. E.	.19								

The correlation values are all positive except two in relearning, four in recognition and one in reconstruction, and the chances are at least 4.5 to 1 that the data will be reproduced in another series of experiments with as few as ten subjects.

The negative correlation between learning ability and the amount of retention as measured by the "saving method" is an obvious consequence of the method of computation. Let X and

Y represent the number of presentations involved in learning and relearning respectively. Then X and $\frac{c}{X}$ will represent deficiency and proficiency of learning ability respectively, and $\frac{X-Y}{X}$ or $1-\frac{Y}{X}$ will represent the percentage value of retention as measured by the "saving method," in which the fraction $\frac{Y}{X}$ will have a value less than 1. The negative correlation between $\frac{c}{X}$ and $1-\frac{Y}{X}$ which we secured means a positive correlation between X and $1-\frac{Y}{X}$ and this is necessarily due to a negative correlation between X and $\frac{Y}{X}$, for as X increases, the value of $1-\frac{Y}{X}$ can increase only as the value of the fraction $\frac{Y}{X}$ decreases.

This negative correlation between X and $\frac{Y}{X}$ will naturally occur because of the presence of X in both values whenever the following relations between X and Y obtain:

1. Absence of correlation.
2. A negative correlation.
3. A positive correlation when $\frac{\text{increment } Y}{Y}$ is less than $\frac{\text{increment } X}{X}$.

In other words, the correlation between $\frac{c}{X}$ and $1-\frac{Y}{X}$ is necessarily negative unless X and Y are positively correlated and at the same time fulfill the condition that $\frac{\text{increment Y}}{\text{Y}}$ is equal to or greater than $\frac{\text{increment X}}{\text{X}}$.

The actual correlation between X and Y determined in these experiments is positive, as presented in Table XXX. The required relation between the increments of X and Y cannot be determined very easily. However, considering the phenomenon of diminished returns in the learning curve, such a relation between X and Y is very unlikely. There is more material to be mastered in the original learning than in relearning. Hence, the increments of X due to deficiency in learning ability will as a rule be proportionally as well as absolutely greater than the respective increments of Y.

Thus, the negative correlation which we ascertained between learning ability and the amount of "saving" is a consequence of the fact that X is involved in both values. It is merely a product of a negative correlation between X and $\frac{1}{\text{X}}$, and a positive correlation between X and Y which we secured. A negative correlation will also obtain under almost every possible relation between X and Y. The correlation thus has no significance as to learning ability and retention.

This analysis confirms our previous contentions that the relearning method constitutes a poor measure of retention.

TABLE XXX

CORRELATION BETWEEN THE SPEED OF LEARNING AND OF RELEARNING

	20 m.	1 h.	4 h.	1 d.	2 d.
Correlation	.50	.21	.98	.35	.78
P. E.	.17	.21	.01	.20	.09

With the exception of relearning, the remaining data may be easily summarized as follows:

(1) With the recognition method and possibly also reconstruction, the correlation between the speed of learning and the amount of retention tends to change from positive to negative

as the interval is lengthened. The fast learners tend to be the least efficient in recognition for the comparatively longer intervals.

(2) With the written reproduction method, the correlation between the same factors is always +, if there is any correlation at all. The fastest learners are as a rule the best retainers.

(1) and (2) together might suggest that the two memory processes differ in quality as well as in quantity, but our data are not conclusive.

(3) While the increase in the degree of learning tends to equalize individual differences, it does not at the outset minimize the numerical value of certain correlations. On the contrary, up to 150% learning, increase in learning seems to make the correlation between learning and retention more definite and reliable.

It seems to the present writer that a more careful study of the correlation between the above two factors is the only systematic way to approach the difference between "immediate" and "permanent" memory. Sometimes it has been asserted that recall immediately after learning differs from delayed recall in nature as well as in quantity.[2] One forgets that immediate recall is recall after a neglected interval. It is neglected because it is too short to be appreciable. The control of this interval will determine a point on the curve of retention just as any other point. Besides certain questionable introspective conclusions, the chief reason for differentiating "immediate" and "permanent" retention in that way is the negative correlation that has occasionally been discovered between the speed of learning and the amount of retention. The relearning method was often used for this purpose and, for some unknown reason or another, one particular time interval was chosen. Now we have discovered that the numerical value of the correlation for the same group of subjects changes under various conditions. It varies with the method of measurement, the degree of learning, the length of the time interval, and perhaps with every variable condition that we can mention. Similar variations will most probably be found in the correlation values for "permanent" memory measured for two

[2] E. g., E. Meumann, "The Psychology of Learning,, Eng. Tr., pp. 40 ff.

different intervals. The objective reason for differentiating "immediate" and "permanent" memory is, therefore, unsound.

5. Correlation between the Speed of Learning and the Speed of Recall.

The data are presented in Table XXXI. Like Table XXX, it contains only the correlation values which are at least twice as high as their respective P. E.'s.

TABLE XXXI

CORRELATION BETWEEN THE SPEED OF LEARNING AND THE SPEED OF RECALL

	20 m.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 d.	2 d.
Writ. Reprod.									
150% learning					.42				
P. E........					.18				
100 %........	.55	.61							
P. E........	.11	.10							
67%	.70							.61	.40
P.E........	.11							.14	.19
33%	.51				.43			.69	.49
P. E........	.16				.18			.12	.17
Recognition									
150% learning			.60	.42		—.45			—.47
P. E........			.14	.18		.18			.17
100%	.58	.56							
P. E.......	.11	.11							
67%									—.63
P. E........									.13
33%									
P. E........									
Reconstruction									
150% learning			.71	.48	.72	.41			
P. E........			.11	.17	.11	.19			
100%	.36	.34							
P. E........	.14	.14							
67%	.48				—.73				
P. E........	.17				.10				
33%									
P. E........									

The speed of recall for written reproduction is taken to be the number of seconds per unit material. For recognition and reconstruction, it is the total duration of the recall process in number of seconds.

Negative values are again found in recognition for the longer intervals. The slow learners are not only more proficient in recognition memory in the long run, but they also recognize the material much faster. Another solitary negative value appears in reconstruction. Its significance is probably the same as for the negative recognition values. Otherwise, the faster learner also tends to be the faster in recall. This correlation is most probably due to the fact that the faster learners are also more proficient retainers. So we present in the next section the correlation between the amount and the speed of recall.

6. Correlation between the Amount and the Speed of Recall.

TABLE XXXII

CORRELATION BETWEEN THE AMOUNT AND THE SPEED OF RECALL

	20 m.	1 h.	2 h.	3 h.	4 h.	6 h.	12 h.	1 d.	2 d
Writ. Reprod.									
150% learning			.84	.80	.72	.96	.78	.96	.61
P. E........			.06	.08	.11	.02	.09	.02	.14
100%	.60	.71			.73			.76	.50
P. E........	.10	.08			.07			.06	.12
67%	.65	.89			.97			.62	.60
P. E........	.13	.05			.01			.14	.14
35%	.65	.82						.80	.89
P. E........	.13	.07						.08	.05
Recognition									
150% learning			.77	.60		.81		.50	.44
P. E........			.09	.14		.08		.17	.18
100%	.63	.47			.59			.34	.62
P. E........	.10	.13			.11			.14	.10
67%	.62	.76			.64			.43	
P. E........	.14	.09			.13			.18	
33%		.38			.43				
P. E........		.19			.18				
Reconstruction									
150% learning			.73	.45		.61	.42		
P. E........			.10	.17		.14	.18		
100%	.72	.71			.42			.54	
P. E........	.08	.08			.13			.12	
67%									—.47
P. E........									.17
33%					.53				.45
P. E........					.16				.17

The individual who retains the most also tends to recall it the most readily. With the exception of a single negative value in reconstruction, this tendency is independent of the speed of learning and does not seem to vary according to the method of measurement, the degree of learning or the length of the time interval.

VIII. CONCLUSION

1. The curve of retention varies with the method of measurement.

A) The curve as determined by the "saving method" differs both in height and general shape from all the other curves determined by methods which directly measure the amount of retention. For the comparatively shorter intervals, the relearning curve falls more rapidly than does most any curve, but the tendency is reversed for the longer intervals. This difference in the shape of the curves follows as a mathematical necessity from the predetermined differences in the units of measurement.

B) Aside from the relearning curve, the other curves stand in the order ot their numerical values as follows: Recognition is the first, reconstruction the second, written reproduction the third and anticipation the last. Many reasons may be given for these particular numerical differences, but above all the order of the curves is a function of the number of restricting factors involved in the conditions of recall.

C) In spite of the variations in the methods of measurement, the curves are, on the whole, more similar to that of Ebbinghaus than to those of Ballard.

D) In numerical value, our retention curve for relearning approaches most closely to that of Finkenbinder, but our recognition curve is vastly different from that of Strong. The similarity of data depends upon the corresponding similarity of technique.

2. The curve of retention varies with the degree of the original learning. The amount of retention for most intervals increases with the degree of learning.

A) Increase in the degree of learning favors the more difficult methods for the longer intervals.

B) The effect of the increase in the degree of learning upon the amount of retention manifests the phenomenon of diminished returns. This phenomenon is the incidental and almost necessary consequence of the tendency of negative acceleration obviously present in the learning curve for memory.

C) On the whole, the curves of retention for the different degrees of learning still approach more closely to the Ebbinghaus type than to that of Ballard. The recognition curves are, however, far from being logarithmic.

3. Beyond a certain limit, the duration of the recall process has but little effect upon the amount of recall or the curve of retention.

A) The shape of the retention curve for written reproduction is practically determined at the end of the first 2 min. of recall. For recognition 90 sec. of recall is long enough for determining the shape of the curve.

B) With the written reproduction method, the effect of extending the time limit for recall upon the amount of reproduction is negatively accelerative.

4. The amount of error in written reproduction increases with the time interval, but the error curve thus determined manifests no definite relationship to the curve of forgetting which also rises with the time interval. Nevertheless, the former has to be taken account of as a supplementary curve of retention.

5. The speed of recall decreases with the time interval, but the speed curve, too, bears no similarity or causal relation to the curve of forgetting, though it is indicative of the conditions of retention.

6. A) Individual variability increases with the difficulty of the act of recall, but decreases with the frequency and the recency of practice.

B) The speed of learning and the amount of retention are positively correlated.

The results of these experiments prove that the difference between the Ebbinghaus tradition and the type of curve discovered by Ballard is not due to differences (1) in the method of measurement or (2) in the degree of the original learning.

We do not find a higher amount of retention for the 2-day interval than for 1 day or even immediate recall, as did Ballard.

Differing from the Ebbinghaus tradition, our curves are not all logarithmic. Some of the recognition curves do not even manifest the phenomenon of negative acceleration in general.

The curve of retention varies with the conditions of learning and of recall.