Summer 1942

The Curve of Forgetting For Substance Material

Leslie J. Briggs

Fort Hays Kansas State College

Follow this and additional works at: https://scholars.fhsu.edu/theses

Part of the Education Commons

Recommended Citation

https://scholars.fhsu.edu/theses/356

This Thesis (L20) is brought to you for free and open access by the Graduate School at FHSU Scholars Repository. It has been accepted for inclusion in Master's Theses by an authorized administrator of FHSU Scholars Repository.
THE CURVE OF FORGETTING
FOR SUBSTANCE MATERIAL

being

A thesis presented to the Graduate Faculty
of the Fort Hays Kansas State College in
partial fulfillment of the requirements for
the Degree of Master of Science

by

Leslie J. Briggs, B. S. in Educ.
Fort Hays Kansas State College

Date July 22, 1942
Approved: J. B. Reek
Major Professor

Chr. Grad. Council
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title Page</th>
<th>ii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>i</td>
</tr>
<tr>
<td>List of Tables in Text</td>
<td>iv</td>
</tr>
<tr>
<td>List of Figures in Text</td>
<td>v</td>
</tr>
</tbody>
</table>

## I. INTRODUCTION

A. The Nature of the Problem
   1. Scope
   2. Practical Importance of the problem

B. Historical Background of the Problem

C. Review of Experiments Related to the Problem
   1. Retention
   2. Reminiscence

## II. METHOD

A. Subjects

B. Materials
   1. The immediate test
   2. The re-test
   3. Reliability of the test
   4. Validity of the test

C. Procedure
   1. Grouping of the subjects
   2. Administration of the tests
   3. Scoring of the tests
### III. RESULTS

A. Recall ................................. 15
   1. Immediate recall .................. 16
   2. Delayed recall .................... 17
   3. The curve of forgetting ........... 19

B. Variability as a Function of Delayed Recall ............... 22
   1. Variability in immediate recall ... 24
   2. Variability in delayed recall ..... 25

C. Reminiscence as a Function of Delayed Recall .......... 26

D. Sex Differences ....................... 29
   1. Sex differences in recall .......... 29
   2. Sex differences in relative variability ...... 31
   3. Sex differences in reminiscence ... 31

### IV. CONCLUSIONS AND DISCUSSION

A. Summary of Conclusions .................. 32
B. Limitations of results .................. 34
C. Inferences ................................ 35

Bibliography .................................. 36
Appendix A ...................................... 42
Appendix B ...................................... 46
List of Tables in Text

I. Comparison of the Means of the Control Group and the Experimental Group on the Immediate Test and at the 12-week Interval........12

II. Average Immediate and Delayed Recall Scores for all Experimental Groups......................15

III. Frequency Distribution and Measures of Relative Variability for Experimental Groups in Immediate and Delayed Recall..............23

IV. Percentages of Cases Showing Reminiscence, No Change, and Decrease in Amount of Material Recalled from the Immediate Test to the Re-test......................26

V. Mean Scores, by Sex, on Immediate and Delayed Recall, and the Reliability of the Mean Differences.................................29

VI. The Standard Deviations and the Coefficients of Relative Variability by Sex for Immediate and Delayed Recall....................31
List of Figures in Text

1. Forgetting Curves from Several Important Memory Experiments................................. 4
2. Curves of Forgetting for Nonsense Syllables, Factual Material, and Substance Material.....20
3. Smoothed Curve of the Coefficient of Relative Variability for Immediate and Delayed Recall..............................................................25
4. Percentage of Cases Exhibiting Reminiscence, and Percentage of Cases Making at Least the Same Score on the Immediate and Delayed Recall Tests...............................28
I. INTRODUCTION

A. The Nature of the Problem

The purpose of this paper is to describe the method and analyze the results of an experiment on substance memory. The term "substance memory," as used in the present study, refers to the retention of general ideas derived from meaningful reading material in which the facts presented require some generalization and abstraction of meaning on the part of the reader. This type of memory is to be contrasted with "factual," or "verbatim" memory, in which the task is to retain specific, detailed facts stated in the reading material, and with rote memory, which involves the memory for illogical, disassociated, discrete units.

1. Scope

The specific object of this investigation was to measure the substance retention of college students for the content of a rather difficult group of paragraphs on psychological topics, and to plot a curve of forgetting for this type of material.

The questions regarding substance memory which this investigation will attempt to answer may be stated as follows:

a. What is the course of the curve of forgetting for substance material for delayed recall intervals of one to twelve weeks?
b. What is the effect of the various intervals of delayed recall upon the variability of the group?

c. What is the effect of the various intervals of delayed recall upon the percentage of the cases exhibiting reminiscence? 1

d. What sex differences are there in immediate and delayed recall of substance material?

It is not within the scope of this paper to investigate the theoretical explanations of how memory takes place, nor to discuss the relative merits of the various methods of measuring memory and presenting the material.

2. Practical importance of the problem

The problem of how much of the material which people read and temporarily learn is actually retained for any length of time is obviously of importance in education. Teachers have long been on the quest for some method of teaching which will bring about permanent retention on the part of students. The rote learning of the past generations is destined to yield to some new approach to the problem of retaining the material learned. The comparison of forgetting curves for "factual" and "substance" material may offer a hopeful suggestion.

1 Reminiscence has been defined by McGeoch as "the improvement in the recall of incompletely learned material after an interval of time without intervening formal relearning or review."

B. Historical Background of the Problem

The volume of discussion and reports of experimentation found in the literature must not lead the reader to the hasty conclusion that the scientific study of memory has long been in existence. As recently as the nineteenth century, the apparent impossibility of subjecting the "higher mental processes" to experimentation led psychologists to divert their attention to the study of psychophysics. It was not until the publication of Ebbinghaus's *Über das Gedaechtnis*\(^1\) in 1885 that the scientific study of memory was begun. Even this pioneer considered memory to be the linking together of discrete units, and employed nonsense syllables and poetry as his materials.

The next progressive step was made when meaningful prose material was employed in the study of memory. From this point, "meaningful material" was sub-divided for experimental purposes into such classes as "factual," "verbatim," "paraphrase," and "substance" material.

---

C. Review of Experiments Related to the Problem

1. Retention

A glance at forgetting curves derived from several major studies of memory emphasize the fact that the percentage of material retained depends largely upon the type of material used. Figure 1 presents graphically the data from several important memory experiments.

**Figure 1**

Forgetting Curves from Several Important Memory Experiments
a. Ebbinghaus's experiments. The experiments of Ebbinghaus, who used nonsense syllables as material and who employed the savings method in measuring retention, indicate a very large loss in retention, especially for the longer time intervals. His results are shown in Figure 1 by the curve which is now well known as the "Ebbinghaus curve." The rapid initial loss, followed by a more gradual final loss, is typical of the forgetting curve for illogical material.

b. Radosawljewitsch's experiments. Radosawljewitsch found greater amounts of immediate retention than did Ebbinghaus, but the final percentages retained in 30 days are comparable. Radosawljewitsch employed both poetry and nonsense syllables as material in his investigations. The curve in Figure 1 is derived from his data on memory for nonsense syllables.

c. Dietze and Jones' experiment. A slightly different curve is shown by an experiment reported by Dietze and Jones. In their study, the material consisted of factual prose selections rather than meaningless material. It

---


2 loc. cit.

should be noted that with this type of material, although the curve begins lower, the final retention is higher than in the materials used in the experiments of Ebbinghaus and Radosawljewitsch. In other words, the final interval of 100 days for meaningful material shows a larger percentage retained than much shorter intervals for the retention of nonsense syllables. For meaningful material, the curve drops less rapidly and the final retention is much higher.

However, differences in methods of measurement used by the various investigators make a comparison of results difficult.

d. The retention curves derived from the studies of English, Welborn, and Killian present a drastic contrast to the curves of Ebbinghaus, Radosawljewitsch, and Dietze and Jones. The distinguishing factor is in the appearance of reminiscence, and an exceedingly high retention over a long period of time. English and his associates1 used both verbatim and substance items in measuring retention, and extended their time intervals to 71 days.

The validity of the increase in retention of substance items up to the 8 day interval, and the very high degree of retention up to the 71 day interval indicated by English,

---

has raised some questions in the mind of the author of the present paper. It is the purpose of this paper to check this unusual rise in retention for substance material.

2. Reminiscence

Reminiscence was first discovered by Ballard in 1913 in his studies of London school children, and study of the phenomenon was taken up by McGeoch, Williams, and English, Welborn, and Killian. Although reminiscence has been found in the median scores in some studies, it is more likely to occur only in individual cases within the group. McGeoch says, "Reminiscence is disguised in averages of groups," whereas "... the per cent of subjects showing reminiscence and the amount shown are adequate measures of reminiscence." He adds that the effect upon reminiscence of such factors as time intervals, age, sex, and intelligence of the subjects is not yet clearly understood.

In later pages, this paper presents some data on reminiscence in which the variables are sex and delayed recall intervals.

---

4 loc. cit.
II. METHOD

The specific method for measuring substance memory in this investigation, consisted of administering mimeographed copies of rather difficult reading material on psychological topics to five groups of college students. The amount of material learned and immediately recalled was measured by an objective recognition-type test. A re-test was given to one group at a time at the various time intervals. The forgetting curve was drawn on the basis of the average retention scores at each time interval.

A. Subjects

The subjects for the experimental groups consisted of 67 college students enrolled in 4 classes of lower-division psychology courses. The subjects for the control group consisted of 28 students enrolled in a lower-division English class. The majority of the students were freshmen, although a few were upper-division students.

Since some of the students were absent during the re-tests, the total number of subjects upon which the data were based was 89.
B. Materials

1. The immediate test.

The reading materials and test questions used for the experiment were taken, with some modifications, from the Van Wagenen Reading Scales in Educational Psychology. The test was shortened by including only the first 12 paragraphs, and some of the original questions were modified to make them more suitable for the experiment. In the mimeographed test used for the experiment, the paragraphs were presented in order, with the questions pertaining to the paragraph immediately below, so that the paragraphs were separated by the appropriate questions. In all, there were 50 true-false questions covering the 12 paragraphs of reading material presented.

In order to make a comparison with English's results legitimate, care was taken to make each question a true "substance item" and not a "verbatim item." The criterion for a true substance item stated by English was observed as carefully as possible.

"...A V-item uses the exact words of a single sentence; sometimes the sentence is simplified by omission, but never by addition except for some negative in order to have false items. The S-item may use a different sentence structure and a dif-

---

1 Van Wagenen, M. J., Van Wagenen Reading Scales in Educational Psychology. Scale A. Published by the Educational Test Bureau, Inc., Minneapolis, Minnesota. 1929.
Different vocabulary as far as possible, and--
the critical point--can never be answered
on the basis of a single sentence in the
text..."1

2. The re-test

For the re-tests, the same 50 questions were presented
in mimeographed form, with the reading material omitted.
Copies of the immediate test and the re-test appear in
appendix A.

3. Reliability of the test

The reliability of the test was determined by the split-
half correlation method. The results of the immediate test
for the original 95 subjects were used in determining the re-
liability. The odd-and even-numbered items were scored sep-
arately and then correlated, yielding a coefficient of cor-
relation of .575 for one-half of the test. The coefficient
of reliability for the whole test, by the use of the Spearman-
Brown formula,2 was found to be .73.

This degree of reliability was considered adequate for
the purposes of this experiment. In evaluating the suffici-

1 English, H. B., A letter dated January 12, 1942. Ohio
State University, Columbus, Ohio.

2 Garrett, Henry F., Statistics in psychology and educa-
tion. 2nd ed. New York, London, and Toronto, Longmans,
ency of this reliability coefficient, it should be men-
tioned that the variability of scores was small, showing
a rather high degree of homogeneity in the subjects.
Increasing the variability would have automatically raised
the reliability coefficient of the test.

4. Validity of the test

The primary assumption involved in using a test of
this type to measure retention is that the learning and
retention which the tests measured were derived from the
reading material presented. Since the test was designed
originally for college juniors, it was believed that the
material presented in the test was not of a nature usually
directly taught in elementary courses in psychology;
however, the control group (the English class) was given
the test at the same time it was administered to the ex-
perimental groups (psychology classes). The control
group was re-tested at the longest experimental time inter-
val, and the mean scores compared.

The results for both groups tested and re-tested at
the same interval are found in Table I.
Table I

Comparison of the Means and Standard Deviations of the Control Group and the Experimental Group on the Immediate Memory Test and at the 12-week Interval of Delayed Recall

<table>
<thead>
<tr>
<th>Group</th>
<th>Immediate Test</th>
<th>Delayed Recall Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>E-Control</td>
<td>34.42</td>
<td>8.22</td>
</tr>
<tr>
<td>D-Exper.</td>
<td>34.92</td>
<td>4.95</td>
</tr>
</tbody>
</table>

These data indicate that the experimental group actually did not learn or relearn more efficiently than did the control group. Since the retention is better for the control group, it appears that it is safe to assume that the re-tests actually measured retention of the materials presented in the experimental readings, and that the scores for the experimental group were not artificially raised through class instruction.

Since some of the test questions were modified to obtain an approximately equal number of true and false questions, those items which were changed were submitted to the members of the psychology staff for examination. Each staff member marked the questions as being true or false, and the questions upon which there was disagreement were further modified to make them more definitely either true or false. In this way, the phrasing of the questions was checked.
D. Procedure

1. Grouping of the subjects

The subjects for the experiment were divided into 5 groups according to the classes in which they were enrolled. Each group was given the immediate test and one re-test at the interval designated. The time intervals for the delayed recall tests were assigned as follows:

- Group A: 1 week
- Group B: 4 weeks
- Group C: 8 weeks
- Group D: 12 weeks
- Group E: 12 weeks (Control)

2. Administration of the tests

The immediate test was administered by the writer to each group at a regular meeting of the class. No special motivation was employed; the students were told that they were being given the test for experimental purposes, and that they would be re-tested over the material sometime later in the semester. No time limit for taking the test was set, for it was desired that each subject should have sufficient time to re-read the material as often as necessary to answer the questions, thus insuring maximum learning for the immediate test. All subjects finished the
test within the class period.

A copy of the instructions appears in the sample test in the appendix.

The re-tests were given without warning during the regular class periods on the day corresponding to the assigned time interval for the groups. The students were told that they were being re-tested over the material they had read in class earlier in the semester. Enough time was given to permit all subjects to finish the questions.

3. Scoring of the tests

For both the test and re-test, separate answer sheets were used, on which the choice answer was indicated by the subject by blacking-out numbers corresponding to "true" or "false," as directed. The directions for marking were indicated on the black board in addition to being printed on the test sheets in order to insure understanding of the marking system. For accuracy, the papers were scored by number of items answered correctly with the aid of a prepared key. It was considered that the absence of a time limit made statistical correction for chance unnecessary.
III. RESULTS

A. Recall

Table II summarizes the results of the memory tests for all experimental groups. The immediate recall scores represent the average amount of the reading material comprehended by each group; in other words, it shows the amount and percentage of the reading material which originally "got across" to the subjects.

The delayed recall scores indicate the average amount of the material which was retained at the various time intervals. The average per cent remembered is obtained by converting the recall scores into percentages of material retained, thus forming the points through which the curve of forgetting was plotted.

Table II

<table>
<thead>
<tr>
<th>Average Immediate and Delayed Recall Scores</th>
<th>for All Experimental Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recalls Mean Score</td>
<td>% Immed.</td>
</tr>
<tr>
<td>Group</td>
<td>N Interval</td>
</tr>
<tr>
<td>A</td>
<td>17</td>
</tr>
<tr>
<td>B</td>
<td>17</td>
</tr>
<tr>
<td>C</td>
<td>12</td>
</tr>
<tr>
<td>D</td>
<td>24</td>
</tr>
</tbody>
</table>

1 The total number of cases was 70. All cases absent from the re-tests were discarded in calculating the means both for immediate and delayed recall. The reliabilities of the means appear in the appendix.
1. Immediate recall

Table II reveals that the college students who took part in the immediate recall test were able to answer an average of 35.43 items out of 50, or 70.86 per cent. It must be remembered that the scores in "immediate recall" are perhaps more largely influenced by reading skill and general comprehension than by actual ability to remember. At any rate, the figure of 70.86 per cent represents the amount of material learned or "gotten across" as a result of the reading, and is termed "immediate recall" rather than "initial comprehension" for the sake of convenience.

The percentage cited above for immediate recall of the "substance" of the material is considerably higher than the percentage of "factual material" immediately recalled in Dietze's experiment with college students. His subjects recalled an average of 58 per cent of the material in an immediate test. (See Figure 1)

On the other hand, other investigations have found the immediate recall for factual items higher than for substance items. In the experiments of English, Welborn and Killian, there is a definite superiority for factual items in the immediate test and up to an interval of 8 days, after which summary items yield better retention. (See Figure 1) However, in this study, the figures show continued reminiscence
for summary items up to 8 days, which indicates that there must have been factors operating which were absent in the present study, and in most others.

The large difference in percentage of substance items immediately recalled between the present study and that of English and others, can probably be explained by noting the difference in procedure. In the present study, the subjects were allowed to re-read any portions of the material necessary to answer the questions, and were permitted to refer back to the material while taking the test. In English's study, the subjects were not permitted to see the questions before reading the material, and they did not have access to the material while answering the questions.

2. Delayed recall

In this paper, the term "delayed recall" is to be interpreted not as the recall following a mere reading of the material, but as the recall following the reading and one immediate test. An immediate test was given all groups in order to determine the similarity in efficiency in learning. Table II indicates that the groups were very similar in learning ability, for the widest difference in immediate recall between any two of the four groups was 1.96 per cent, or .98 of a single score. This result made pairing of scores unnecessary, for it appears that the chance distribution of the students into the various groups was as favor-
able for experimentation as an artificial parallel grouping would have been. Therefore, the comparison of average scores for the groups re-tested at various time intervals was deemed legitimate.

Table II reveals delayed memory to be 68.58 per cent after one week, and 63.42 per cent after four weeks. There is another drop to 61.16 per cent for the 8-week interval, but from 8 to 12 weeks there is no drop; in fact, the figures show a very slight gain, which however, is entirely within the standard error of the means. The reliabilities of the differences between immediate test scores and delayed test scores are found in the appendix.

It was the original plan of this experiment to include a 15-week interval of delayed memory. Group F, consisting of 22 subjects, was given the immediate recall test, yielding a mean score of 36.36. However, just before the group was re-tested at the 15-week interval, they were unintentionally given instruction in class on topics covered by the experimental material. This instruction artificially influenced the re-test, which yielded a mean score of 32.86. Since this result is significantly above the re-test score for the 12-week interval, the only possible explanation lies in the effect of the class instruction.

However, in order to verify this conclusion, Group B, which had taken the re-test at the 4-week interval, was
again re-tested at a 16-week interval, with a resulting mean score of 31.55. Disregarding the undisputed advantage of the practice afforded by two previous tests, this score is still lower than the obtained score of Group F at the 15-week interval, demonstrating the impossibility of the Group-F score being accomplished without the supplementary instruction in class. Therefore, it was necessary to exclude the 15-week delayed recall score from the results in Table II, and from any further mention in this paper.

Proof that there was no such artificial rise in scores up to the 12-week interval has already been given under the section on validity of the test.

3. The curve of forgetting

The following curve of forgetting for substance material is constructed on the basis of the data given in Table II.

The dot at the beginning of the curve represents the average of all groups on the immediate recall tests; the other dots represent delayed recall for each group re-tested at the various time intervals.

Ebbinghaus's curve for memory of nonsense syllables and Dietze's curve for factual memory are reproduced here for comparison.
It is evident that the curve for substance memory differs radically from the curve presented by English, (Figure 1), who found the score for the 71-day interval to be above the immediate test score.

It is also evident that this curve for substance memory drops less in rapidity and amount than Dietze's curve for factual memory, which in turn, drops less rapidly than Ebbinghaus's curve for memory of nonsense syllables.
The results of McGeoch and McKinney\(^1\) concur with English's results in finding reminiscence for substance memory and definite losses for verbatim memory.

Austin,\(^2\) on the other hand, considered her curve for substance memory similar to the Ebbinghaus curve. Amid this conflicting evidence, one may be tempted to conclude with Carr\(^3\) that there are as many curves as there are variables, or with Davis and Moore,\(^4\) that there is no difference between the curve for nonsense memory and the curve for meaningful material. However, if the varied conditions under which all these experiments were conducted are considered, there is at least a discernable difference between curves of forgetting for rote and logical memory.

The further question of whether there is a qualitative difference between the two types of meaningful material in memory, (viz., substance and verbatim, as is suggested by English), needs more clarification, but the present study indicates superior retention for substance material, though not to the extreme degree found by English.

To sum up the data reviewed above, the author's con-

---

2 ibid., p. 3.
3 loc. cit.
4 loc. cit.
Conclusion is well stated in the words of Reed, who says:

"...Memory for prose substance is thought by some psychologists to be totally different from memory for rote learning, but it probably differs from it in degree rather than in kind." However, "...there is little doubt that the amount retained for meaningful prose is far greater than for meaningless material."

The results presented in the present paper indicate that the category of meaningful material, when divided further into substance and factual material, yields a quantitative difference in retention in favor of the substance material.

B. Variability as a Function of Delayed Recall

Investigators in the field of memory have placed little emphasis on variability in memory as a function of delayed recall. However, an examination of Dietze's tables in the appendix of his thesis, reveals that the variability of groups at an interval of 100 days was much smaller than the variability on the immediate tests.

The following table summarizes the results of the present study on variability as a function of delayed recall.

---


2 Dietze, Alfred G., Factual memory of secondary school pupils for a short article which they read a single time. (Doctor's thesis, University of Pittsburgh, 1930. Appendix B.)
Table III

Frequency Distribution and Measures of Relative Variability for Experimental Groups in Immediate and Delayed Recall

<table>
<thead>
<tr>
<th>Score</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
</tr>
</thead>
<tbody>
<tr>
<td>44-45</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>42-43</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>40-41</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>38-39</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>36-37</td>
<td>-</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>34-35</td>
<td>1</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>32-33</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>30-31</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>28-29</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>26-27</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>24-25</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>22-23</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20-21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>17</th>
<th>17</th>
<th>12</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>35.90</td>
<td>35.82</td>
<td>35.08</td>
<td>34.92</td>
</tr>
<tr>
<td>SD</td>
<td>5.34</td>
<td>3.75</td>
<td>5.87</td>
<td>4.95</td>
</tr>
<tr>
<td>SEM</td>
<td>1.33</td>
<td>3.94</td>
<td>1.77</td>
<td>1.03</td>
</tr>
<tr>
<td>V</td>
<td>14.87</td>
<td>10.47</td>
<td>16.73</td>
<td>14.18</td>
</tr>
</tbody>
</table>

Delayed Recall

<table>
<thead>
<tr>
<th>Score</th>
<th>1 week</th>
<th>4 weeks</th>
<th>8 weeks</th>
<th>12 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>44-45</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>42-43</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>40-41</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>38-39</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>36-37</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>34-35</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>32-33</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>30-31</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>28-29</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>26-27</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>24-25</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22-23</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20-21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>17</th>
<th>17</th>
<th>12</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>34.29</td>
<td>31.71</td>
<td>30.58</td>
<td>30.75</td>
</tr>
<tr>
<td>SD</td>
<td>4.17</td>
<td>4.54</td>
<td>4.19</td>
<td>2.63</td>
</tr>
<tr>
<td>SEM</td>
<td>1.04</td>
<td>1.14</td>
<td>1.26</td>
<td>0.55</td>
</tr>
<tr>
<td>V</td>
<td>12.16</td>
<td>14.32</td>
<td>13.70</td>
<td>8.55</td>
</tr>
</tbody>
</table>
In reference to Table III, it should be stated that since the number of cases was small, due to large class absences, all statistical measures were computed from formulae adapted to individual scores rather than to scores in a frequency distribution.\(^1\)

1. Variability in immediate recall

In comparison of variabilities, it will be noticed that the coefficient of relative variability, \( V \), was used rather than the SD, in order to take into account the differences between the means,\(^2\) which is slight in the case of the four groups in immediate recall, but which is large in the figures for delayed recall.

The coefficients of relative variability indicate that the four groups were fairly similar in variability for the immediate recall tests. Group B shows the greatest deviation in this respect, and it also does not conform to the general trend of the rest of the groups in that there is more variability in this group on the immediate test than on the re-test. The average \( V \) for the four groups on the immediate test is 14.06.

---


\(^2\) *ibid.*, p. 52.
2. Variability in delayed recall

It is evident that in delayed recall, the variability, with some exceptions, tends to become smaller as the time interval becomes longer. This is especially true for the 12-week interval, which shows a very marked decrease in variability.

For convenience, a smoothed curve of relative variability is presented below. The immediate variability is taken as the average $V$ of the four groups on the immediate test.

![Figure 3](image)

Figure 3

Smoothed Curve of the Coefficient of Relative Variability for Immediate and Delayed Recall

Although the number of cases is too small to justify any clear-cut conclusions, the indication is that relative variability decreases as the interval of delayed recall is length-
ened. In other words, there seems to be a tendency for the scores to regress toward the mean as the time interval is lengthened. This is especially true for the 12-week interval, although the wide deviations of each point from the smoothed curve do not permit specific conclusions as to the amount of direct functional relationship between relative variability and length of time interval.

C. Reminiscence as a Function of Delayed Recall

Taking a cue from English and McGeoch, the percentage of cases showing reminiscence for each delayed recall interval were calculated. Table IV summarizes the evidence of reminiscence found in this investigation.

Table IV

Percentages of Cases Showing Reminiscence, No Change, and Decrease in Amount of Material Recalled from the Immediate Test to the Re-Test

<table>
<thead>
<tr>
<th>Score Change</th>
<th>Group A 1 week</th>
<th>Group B 4 weeks</th>
<th>Group C 8 weeks</th>
<th>Group D 12 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reminisc.</td>
<td>29.2</td>
<td>17.6</td>
<td>25.0</td>
<td>12.5</td>
</tr>
<tr>
<td>No Change</td>
<td>11.80</td>
<td>5.9</td>
<td>.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Decrease</td>
<td>59.0</td>
<td>76.4</td>
<td>75.0</td>
<td>75.0</td>
</tr>
</tbody>
</table>

Again, these figures disagree with the findings of English. For substance material, he found reminiscence in the average scores of the groups up to an interval of 8 days, and at the end of 71 days the average was still above the immediate scores.
(See Figure 1) In a later study, he found reminiscence increasing from 58 per cent to 69 per cent in 30 days. Both the amount and duration of reminiscence are greater in the study just cited than in his experiment referred to in Figure 1.

McGeoch has found reminiscence in the analysis of individual scores where average scores concealed it. King and Homan, and Welborn and Killian, have also found evidences of reminiscence in individual scores.

The results of the present study concur with the latter investigators in finding an increase in retention in individual scores, but it differs with English in his reminiscence for median scores.

The percentage of cases exhibiting reminiscence in the present study is shown in a smoothed curve (Figure 4), the points being plotted from the results in Table IV.

In order to include all the subjects who did not show a loss in retention from the immediate to the delayed recall test, Figure 4 includes another smoothed curve which includes both the percentage of cases showing reminiscence and the percentage of cases making the same score on the immediate test

---


4 loc. cit.
and the delayed recall test.

Figure 4

Percentage of Cases Exhibiting Reminiscence, and Percentage of Cases Making at least the Same Score on the Immediate and Delayed Recall Tests

At no time interval did the percentage of cases exhibiting reminiscence reach 50 per cent, which would be the lowest per cent which could cause an increase in median scores. This is true also when the scores remaining the same are counted as showing reminiscence.

Although the present study does not indicate the percentage of cases showing reminiscence for intervals shorter than one week, the data indicate that no such marked degree of reminiscence as found by English would have been discovered for the shorter intervals.
D. Sex Differences

Because of the small number of cases, it was considered inadvisable to separate the scores of each experimental group by sex for the study of sex differences. However, the scores were separated according to sex for all students taking the immediate recall test. This was also done for both groups taking the 12-week delayed recall test, viz., the control group E and the experimental group D. The scores for both groups were combined according to sex, and the differences in mean scores and variability were calculated.

1. Sex differences in recall

Table V summarizes the data on sex differences in recall.

Table V

Mean Scores, by Sex, on Immediate and Delayed Recall, and the Reliability of the Mean Differences

<table>
<thead>
<tr>
<th></th>
<th>Immediate Recall All Groups</th>
<th></th>
<th>Delayed Recall-12 weeks Groups D &amp; E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
</tr>
<tr>
<td>Cases</td>
<td>64</td>
<td>84</td>
<td>22</td>
</tr>
<tr>
<td>Mean</td>
<td>34.94</td>
<td>36.26</td>
<td>30.23</td>
</tr>
<tr>
<td>SEM</td>
<td>.71</td>
<td>.60</td>
<td>.57</td>
</tr>
<tr>
<td>SEDiff.</td>
<td>.73</td>
<td></td>
<td>1.15</td>
</tr>
<tr>
<td>MD</td>
<td>1.30</td>
<td></td>
<td>1.58</td>
</tr>
<tr>
<td>D/SEDiff.</td>
<td>1.4</td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td>Chances in 100 of true diff.</td>
<td>92</td>
<td></td>
<td>92</td>
</tr>
</tbody>
</table>
a. Immediate recall. Table V indicates a mean score difference in immediate recall of 1.3, or 2.6 per cent in favor of the girls. The Critical Ratio of 1.4 does not prove absolute reliability, but the difference found is at least suggestive.

b. Delayed recall. It appears that the girls are also superior to the boys in delayed recall of substance material at the interval of 12-weeks, the difference being slightly larger than in the case of immediate recall. The Critical Ratio is 1.4 here, also. Although the difference between the means was not proved to be significant, the results agree with the findings of Wissler, Gates, Lodge and Jackson, and Travis in their agreement that "women are superior to men in substance memory."1

In factual memory, Dietze found that for pupils in grades 7 to 12, the boys were slightly superior to girls, but the difference was explained as being due to differences in interest in the reading material.2 This sex difference for factual material may or may not be true for college students.

---

2. Sex differences in relative variability

Table VI presents the results of this investigation for sex differences in variability in immediate and delayed recall.

Table VI

The Standard Deviations and the Coefficients of Relative Variability by Sex for Immediate and Delayed Recall

<table>
<thead>
<tr>
<th></th>
<th>Immediate Recall</th>
<th></th>
<th>Delayed Recall-12 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Groups</td>
<td>Groups D &amp; E</td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number (N)</td>
<td>64.</td>
<td>22.</td>
<td>21.</td>
</tr>
<tr>
<td>Standard Deviation (SD)</td>
<td>5.66</td>
<td>5.49</td>
<td>2.60</td>
</tr>
<tr>
<td>Variance (V)</td>
<td>16.20</td>
<td>15.15</td>
<td>8.60</td>
</tr>
</tbody>
</table>

These results indicate that the boys are slightly more variable in immediate recall, while the girls are much more variable in delayed recall for the 12-week interval.

3. Sex differences in reminiscence

The percentage of cases showing reminiscence at the 12-week interval was found to be 13.63 per cent for boys, and 14.28 per cent for girls. These figures agree with the average retention scores, which were also higher for the girls than for the boys.
IV. CONCLUSIONS AND DISCUSSION

A. Summary of Conclusions

The following conclusions are presented as resulting from the data and discussion in previous sections:

1. The amount of immediate recall of college students for the "substance" of rather difficult prose selections on psychological topics was found to be 70.86 per cent.

2. For delayed recall, the average percentage of "substance" retained was found to be 68.58 per cent at 1-week, 63.42 per cent at 4-weeks, 61.16 per cent at 8-weeks, and 61.50 per cent at 12-weeks. The curve of forgetting for substance material derived from this experiment was found to drop much less than curves of forgetting for nonsense and factual material, as drawn by other investigators. However, the very great superiority of substance memory over verbatim memory at long time intervals in terms of reminiscence as found by English, was not verified by this experiment. The difference between substance memory and verbatim memory is probably actually a difference in degree rather than in kind.

3. Variability of groups in delayed recall for substance material seems to decrease as the length of time interval increases, although the trend was not consistent enough to justify a statement of an exact functional relationship.
4. The percentage of cases exhibiting reminiscence in delayed recall for substance material decreases with the lengthening of the delayed recall interval, and describes a curve similar to the curve of forgetting for substance material.

5. For college students, girls were found to be superior to boys in both immediate and delayed recall for substance material, although the difference is not completely reliable.

6. College boys were found to be more variable than college girls in immediate recall for substance material, but the positions were reversed for delayed recall at the 12-week interval.

7. The percentages of cases showing reminiscence at the 12-week interval of delayed recall was larger for girls than for boys.
B. Limitations of Results

The reader should remember that the degree of retention of substance material indicated by the present study is conditioned by the method and procedure employed. This is necessarily true for all memory experiments, and it is this factor which causes much of the disagreement of results among investigators.

The relative success in keeping all test items true "substance items" may be another important source of differences in results.

The present investigation encountered one other difficulty—that of obtaining sufficiently large samples for each delayed recall test. This was due to the relatively high incidence of class absences in the experimental groups on the particular days the re-tests were given.
C. Inferences

Although the curve of retention of substance material derived from this experiment differs widely from that of English, it is significant that both curves show much better retention for substance material than do curves for verbatim material. This indicates that in the educational situation, substance material should be emphasized if good retention is to be expected of the students. In other words, more retention may be expected in material which involves general ideas than in material consisting largely of specific facts.

The problem of substance memory offers a fertile field for future research, for probably the question of retention in general will not be answered fully until the various types of materials used are further sub-divided into many specific memory tasks. Perhaps an analysis of these many memory abilities will even reveal the "factors" which cause the difference in performance between "substance" and "verbatim" memory.
Bibliography


A discussion of various factors which are correlated with retention of substance ideas.


Results and discussion of experimental work in the field of psychology.


An outline of the progress of experimental methods in the investigation of psychological problems.


A textbook covering the conditions and theories of learning and retention.
Dietze, Alfred G.  *Factual memory of secondary school pupils for a short article which they read a single time.* (Doctor's Thesis, University of Pittsburgh, 1930).

An extensive experiment on factual memory, including correlation studies.


A summary of the main author's Doctor's Thesis on factual memory.


A report of the incidence of reminiscence in an experimental study, and a suggestion for the measurement of reminiscence.

English, Horace B.  Columbus, O., Jan., 12, 1942.

A letter stating a criterion for distinguishing between substance and verbatim test items.

A summary of various experiments in substance memory, with the conclusions on the basis of the combined work of the three authors.

Garrett, Henry E.  *Statistics in psychology and education.*  

A guide for the statistical treatment of experimental data.

Gates, A. I.  "Correlation of immediate and delayed recall."  

A refutation of Neumann's theory that those who learn quickly forget quickly.


A demonstration of the fact that logical material is retained better than illogical material.

A discussion of the advantages of substance learning over rote and nonsense learning, with comments on the methods of recall and recognition in studying memory.


A definition of reminiscence, with comments on the merit of this phenomena to more attention from experimental psychology.


A demonstration of retroactive inhibition as a major condition in forgetting.


An experiment demonstrating the relative amount of retroactive inhibition in prose and nonsense learning and memory.

A comparison of the amounts of forgetting as a function of the difficulty of the learning material and the ability of the subjects.


A review of experimental findings in learning and memory, with emphasis on the difference in retention of prose and rote material.


A discussion of the importance of meaning in learning material as a partial substitute for routine drill in learning.


An experiment in the retention of nonsense syllables over various time intervals.
Van Wagener, M. J. *Van Wagener reading scales in educational psychology*. Scale A. Published by the Educational Test Bureau, Inc., Minneapolis, Minn. 1929.

A reading test designed to measure level of comprehension of college juniors in educational psychology.


A review of memory experiments with meaningful material, with a short history of the problem and a bibliography of studies relating to memory.


A review of experimental findings on retention, including theories of retention.
Appendix A

Exhibits of Materials
SELECTED READINGS IN PSYCHOLOGY
Recognition Test
Experimental Study in Substance Memory

DIRECTIONS: Read the first paragraph through carefully. Then read the first statement below the paragraph. If the idea it expresses is directly stated in the paragraph or if it can be rightfully inferred from the whole paragraph, draw a circle around the small 1 on the first line of the yellow check sheet.

If the idea expressed in the statement is false—i.e., if it is not stated in the paragraph or cannot be inferred from it, draw a circle around the small 2 on line one of the check sheet.

Do the same for all the questions in the paragraph, and then go on with the next paragraph.

You may re-read each paragraph as often as necessary.

Time—45 minutes.

A. It has been customary for teachers to regard with suspicion the child who learns his lesson in very much less time than the rest of the class need. The maxim "Easy come, easy go" has been firmly fixed with respect to memory work. Recent experiments with both children and adults as subjects prove conclusively that the quick learner is not the quick forgetter. Children who learn quickly retain more on the average than those who learn slowly, both as tested by immediate and permanent memory. It is very important that all those dealing with children bear this fact in mind. The quick learner, whose work is looked upon with suspicion, and so is sent back to it again and again, is not only developing an emotional attitude of dislike or indifference for the subject and sometimes even for the school, but he is forming bad habits of work. He is learning not to put his best work into his study, not to work at his highest speed, because it "doesn't pay." He forms habits of half-hearted work, of divided attention and the teacher is to blame. Many children of bright minds and quick memories may thus have been almost ruined for their best work, just because their ability was not given full rating, was not accepted at face value. Of course, all children have to be taught to test themselves when they are studying, and to know when they do know the lesson, and not to stop just short of the threshold of recall but rather to go a little beyond. All children need this training, but the quick learner does not need it any more than the slow learner.

1. Children who learn quickly forget readily.
2. Requiring children to study what they have already well learned may develop at least two very undesirable traits.
3. By requiring bright children to work at their highest speed an emotional attitude of dislike or indifference for the subject is developed.
4. The Maxim, "Easy come, easy go," has been conclusively disproved by experiments.

B. A third unwise tendency is to degrade the mere giving of information—to belittle the value of facts acquired in any other way than in the course of deliberate effort by the pupil to relieve a problem or conflict or difficulty. As a protest against merely verbal knowledge, and merely memoriter knowledge, this tendency to belittle mere facts has been healthy, but as a general doctrine it is itself equally one-sided. Mere facts not got by the pupil's thinking are often of enormous value. They may stimulate to active thinking just as truly as they may stimulate to the reception of facts.

A fourth false inference is that whatever teaching makes the pupil face a question and think out its answer is thereby justified. This is not necessarily so unless the question is a worthy one and the answer that is thought out an intrinsically valuable one and the process of thinking used one that is appropriate for that pupil for that question. Merely to think may be of little value. To rely much on formal discipline is just as pernicious here as elsewhere. The tendency to emphasize the methods of learning arithmetic at the
C. The name of Binet will always be prominent in any discussion of intelligence testing. He at first tried to find some experimental test that would be a reliable index to a child's mentality. He measured the lung capacity of boys, their height and weight; he made a study of the food they ate, measured their heads, and even investigated their handwriting with a view toward finding a measure of intelligence. Finally he came to the conclusion that intelligence shows itself in a general way in what one has accomplished in the situations that are common to all. He therefore organized a series of short, heterogeneous tests and arranged them in the order of increasing difficulty. A measure of the difficulty of the tests was arrived at by actually giving them to children and noting their reactions. He thus organized his tests in a scale form. When a series of tests were successfully performed by from 60 to 80 per cent of the children of a given age they were considered as the normal tests for the children of that age.

9. Binet's final tests were based on his conclusion that intelligence is reflected in the responses made to common situations.

10. A test was considered difficult enough and appropriate for any age by Binet when from two-thirds to three-fourths of children of that age could do it.

11. Binet at first tried to find a reliable index of mentality in measurements of physical characteristics.

12. Binet's final tests were based on the results of his measurements of lung capacity, height, weight, size of head, and quality of handwriting.

D. Of more significance is our thought of all forms of play as distinct from work or drudgery. The difference here is not primarily one of the kind of activity, but one of attitude. No given activity can arbitrarily be placed in either class. Listening to a concert, working problems in mathematics, sewing or painting, attending a reception, playing a game of whist, taking a walk, working in the garden,—any one of these may be work of the hardest kind to one person and the most delightful play to another. This difference in attitude is caused by the difference in certain characteristics of the activity. When the activity is considered as work, it is being engaged in, not for its own sake, but because of some result worth while, only to be reached by means of the given activity. The eye of the worker is fixed outside of the activity on the result beyond. When the activity seems play to the individual, the process itself seems worth while; he is concerned only with the activity, that in itself satisfies him. The same result may be obtained as in the former case, but it is not the most important thing to the one engaged in the activity. When it is work the process is merely a means to a desirable end, but when it is play the two are fused, and the process with its result seems desirable.

13. The difference between work and play lies in the activity being done.

14. Any activity is definitely either work or play for all people.

15. Listening to a concert may involve more drudgery than working problems in mathematics.
16. When an activity is carried on merely to attain some desired end the process is work.

17. In play the activity is always satisfying and desirable in itself.

E. The question of the relative value of various avenues of presentation is usually linked with the problem of mental types. Some of the older books on psychology urged the teacher to discover the type of imagery used by each pupil and present material to harmonize with it. This advice was based on the assumption that children fell into definite image types, as the "eye-minded" type, the "ear-minded" type, the "motor-minded" type, etc.; and on the further assumption that when the imagery is predominantly visual, material should be presented through the eye, and when predominantly auditory, through the ear, and so on.

Both assumptions are wrong. Nearly all children can and do employ all of the common types of imagery, which may vary from time to time and from one kind of task to another. Furthermore, when one type is favored—for example auditory or motor imagery, impressions received through the eye may be at once converted into it.

Vivid imagery, as many studies have abundantly shown, is not essential to effective learning of school subjects. The reactions demanded in spelling, reading, drawing, or history may be acquired without any one variety of imagery. Pupils whose concrete imagery is very vivid are in general as likely to be dull as bright, and successful as unsuccessful, in particular subjects.

18. Teachers are urged by the leading books on psychology to discover the predominant type of imagery used by a child.

19. The fact that the visual type of imagery may predominate in a child's mental life is not an adequate reason for visual presentation of material.

20. The importance of imagery in learning was overemphasized in earlier books on psychology.

21. Vivid imagery of the visual type is essential for success in spelling and in drawing.

F. Real Learning. The significant thing is that the child at birth has the mechanism of the ordinary reflexes, but only vast potentiality for the acquisition of conditioned reflexes. This gives it its great capacity for learning. And it should be noted that real learning, in distinction from the mere performance of activities for which the neural mechanism is congenital, consists largely in the acquisition of such conditioned reflexes. Hough has put this very fittingly in contrasting such acquisitions as what we call learning to walk with the acquisition in learning to talk. The child cannot walk at birth because the neural mechanism which functions walking is not yet fully developed: but there seems good reason to believe that as soon as this neural mechanism is developed, as indicated by such studies as those of Kirkpatrick, the child walks without any real learning.

It is very different with such acquisitions as that of learning to talk. There may be an inherited mechanism, and undoubtedly is, for making articulate sounds; but the learning to speak a particular language is a case of the acquisition of a vast number of conditioned reflexes.

22. Walking may be acquired without learning.

23. Children cannot speak their native language at birth because there is no nervous mechanism for it.

24. Children learn to talk without any real learning, just waiting for the neural mechanism to mature.

25. For the ordinary reflexes the neural mechanism is inherited.
G. To the probable fact that original nature is imperfectly adapted to our environment, special emphasis should be given, because in popular belief and in not a few books on pedagogy the doctrine of nature's infallibility has been upheld. It is asserted that all instincts must have some utility, perhaps not always perceptible to man. Instincts, it is said, exist to perpetuate the individual, or the race, and in particular they serve to avoid danger, secure food, and so on. All of this is true only in a general way. On the whole, instinctive capacities do have a utility, particularly under more primitive conditions of life, but the exceptional instances are many and important. Our native equipment provides only a rough adjustment to the environment; just good enough, so that with the assistance of the capacity to learn, the species as a whole manages to survive. Many species have not survived and that those that have, still experience difficulties in adjustment, are proof enough of the fallibility of our inherited equipment.

26. Some books on pedagogy uphold the idea that there is a perfect adaption of instincts to environment.

27. The infallibility of nature implies that all instincts are useful even though their use is not apparent.

20. Instinctive adjustments have been too imperfect to enable all species to survive.

29. The evidence that original nature is only imperfectly adapted to our environment is still insufficient.

II. A distribution is not sufficiently described by stating merely its average or central tendency. A second important fact which is needed is a summary statement of the amount of dispersion around the average. This dispersion or variability, also called variation, spread, scatter, fluctuation, and deviation, may be expressed in terms of any one of several measures. The fairly common ones are the range, sometimes called the absolute range, the quartile deviation or semi-interquartile range, the 10-90 percentile range, the mean deviation, the standard deviation, and the median deviation, which is usually but inappropriately called the probable error. It will be recalled that an average is a point on the scale about which the measures cluster. A measure of variability, on the other hand, is a distance on the scale. This distance is measured from an average and includes a definite proportion of all the cases. Not infrequently a measure of variability is used as a unit distance in terms of which to express the distance of any particular measure from another measure or from an average. For example, one may speak of one point being 2 mean deviations, or 3.4 standard deviations from another.

Most uses of measures of variability are based on the assumption of a symmetrical distribution of scores, and unless this assumption is true some error is present.

30. The average of any distribution of scores is a distance on the scale.

31. Both an average and a measure of variability are essential in adequately describing a distribution.

32. The standard deviation is a measure of variability.

33. A distribution is adequately described by stating some measure of its variability.

I. By the question "what sort of knowledge after all really is furnished by one's senses," we are straightway brought upon a fact which, save for people having become callous to it through life-long habituation, would certainly excite universal amazement. For, as depicted by all the natural sciences, including particularly for our purposes physics and physiology, the characters perceived by means of the senses, far from being any part or parcel of the material things themselves, are not even directly connected with these. Thus, the apparent gauziness of the tree depends upon certain movements of quite another thing, namely, the ether in contact with the percipient's eye. The matter of the external world—so the sciences unanimously teach—is really everywhere in a perpetual state of most violent commotion. It is as much so in the seemingly slumberous surface of a lake as in the storm dashed foam.
It is a sound assumption that observation is an active process. Observation is exploration, inquiry for the sake of discovering something previously hidden and unknown, this something being needed in order to reach any end, practical or theoretical. Observation is to be distinguished from recognition, or perception of what is familiar. The identification of something already understood is, indeed, an indispensable function of further investigation; but it is relatively automatic and passive, while observation proper is searching and deliberate. Perception refers to the already mastered; observation is concerned with mastering the unknown. The common notions that perception is like writing on a blank piece of paper, or like impressing an image on the mind as a seal is imprinted on wax or as a picture is formed on a photographic plate (notions that have played a disastrous role in educational methods), arise from a failure to distinguish automatic recognition and the searching attitude of genuine observation.

The common idea of perception is false.

Becoming aware of something previously unknown, even without active effort, is attentive observation.

Recognition plays no part in genuine observation.

Perception, as compared to observation, is a more active process.

Effect of Practice.—Can we improve memory by practice? We are constantly having experience, and consequently have constant practice in retention. We have as much practice in retention as we do in sensation. Sensation does not need, in ordinary cases, special exercise to develop it. Will special exercise in memorizing permanently improve the retentive capacity of the brain? James, some thirty years ago, answered this question as follows: "All improvement of the memory lies in the line of elaborating the associates of each of the several things to be remembered. No amount of culture would seem capable of modifying a man's general receptiveness. This is a physiological quality, given once and all with his organization, and which he can never hope to change." The great amount of experimental work in memory gives us no reason for modifying James' statement. The experiments do prove beyond doubt, however, that we can greatly improve our ability to memorize. In all kinds of material at least some improvement comes through practice, from the learning of nonsense syllables on the one hand to the learning of the logical material of a book on the other. But that this improvement is in any sense due to an improvement of the retentivity of the brain it would be difficult to prove.

Much more practice in sensation than in memory because more is needed in sensation.

Much experimental work has been done in memory.
45. Ability to memorize has a different meaning than brain retentivity.

46. James denied that there could be any improvement in memory.

47. Practice improves both memory of nonsense syllables and of logical learning.

48. Defects of speech and paralysis are usually due to injuries of the brain.

49. Stimulation of the upper parts of the brain causes movements in the upper part of the body.

50. All experiments in localization have been performed on animals.
Di rection s:

This is a retest over the material you read in this class on February 4 or February 4. In the following statements, black out the small 1 on the correct line of the yellow answer sheet if you remember the statement as being true. Black out number 2 on the proper line of the check sheet if you remember the statement as being false.

1. Children who learn quickly forget readily.
2. Requiring children to study what they have already well learned may develop at least two very undesirable traits.
3. By requiring bright children to work at their highest speed an emotional attitude of dislike or indifference for the subject is developed.
4. The Maxim, "Easy come, easy go," has been conclusively disproved by experiments.
5. Facts obtained in other ways than by thinking may stimulate pupils to active thinking.
6. No teaching that makes the pupils think out answers to questions should meet with approval.
7. Memoriter and verbal knowledge has no value and hence should be excluded from education.
8. All teaching that makes the pupils think out answers to questions is commendable.
9. Binet's final tests were based on his conclusion that intelligence is reflected in the responses made to common situations.
10. A test was considered difficult enough and appropriate for any age by Binet when from two-thirds to three-fourths of children of that age could do it.
11. Binet at first tried to find a reliable index of mentality in measurements of physical characteristics.
12. Binet's final tests were based on the results of his measurements of lung capacity, height, weight, size of head, and quality of handwriting.
13. The difference between work and play lies in the activity being done.
14. Any activity is definitely either work or play for all people.
15. Listening to a concert may involve more drudgery than working problems in mathematics.
16. When an activity is carried on merely to attain some desired end the process is work.
17. In play the activity is always satisfying and desirable in itself.
18. Teachers are urged by the leading books on psychology to discover the predominant type of imagery used by a child.
19. The fact that the visual type of imagery may predominate in a child's mental life is not an adequate reason for visual presentation of material.
20. The importance of imagery in learning was overemphasized in earlier books on psychology.
21. Vivid imagery of the visual type is essential for success in spelling and in drawing.
22. Walking may be acquired without learning.
24. Children learn to talk without any real learning, just waiting for the neural mechanism to mature.

25. For the ordinary reflexes the neural mechanism is inherited.

26. Some books on pedagogy uphold the idea that there is a perfect adaptation of instincts to environment.

27. The infallibility of nature implies that all instincts are useful even though their use is not apparent.

28. Instinctive adjustments have been too imperfect to enable all species to survive.

29. The evidence that original nature is only imperfectly adapted to our environment is still insufficient.

30. The average of any distribution of scores is a distance on the scale.

31. Both an average and a measure of variability are essential in adequately describing a distribution.

32. The standard deviation is a measure of variability.

33. A distribution is adequately described by stating some measure of its variability.

34. Motion plays a relatively small part in the things our senses perceive.

35. It is the movements transmitted from the violent commotion of matter that are directly connected with the characters that we perceive as colors or sounds.

36. It is from the vibratory movements that traverse the ether that visual qualities are perceived.

37. Translatory movements of matter are less rapid than the vibratory movements.

38. The qualities of our perceptions are in the material world around us.

39. The common idea of perception is false.

40. Becoming aware of something previously unknown, even without active effort, is attentive observation.

41. Recognition plays no part in genuine observation.

42. Perception, as compared to observation, is a more active process.

43. We get much more practice in sensation than in memory because more is needed in sensation.

44. Much experimental work has been done in memory.

45. Ability to memorize has a different meaning than brain retentivity.

46. James denied that there could be any improvement in memory.

47. Practice improves both memory of nonsense syllables and of logical learning.

48. Defects of speech and paresis are usually due to injuries of the brain.

49. Stimulation of the upper parts of the brain causes movements in the upper part of the body.

50. All experiments in localization have been performed on animals.
<table>
<thead>
<tr>
<th>Grade</th>
<th>School</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>193...</td>
</tr>
</tbody>
</table>

Match question number in quiz with line number on this score sheet. Encircle number in line which represents your choice of answer.

| 1 | 1 | 2 | 3 | 4 | 5 | 101 | 1 | 2 | 3 | 4 | 5 |
| 2 | 1 | 2 | 3 | 4 | 5 | 102 | 1 | 2 | 3 | 4 | 5 |
| 3 | 1 | 2 | 3 | 4 | 5 | 103 | 1 | 2 | 3 | 4 | 5 |
| 4 | 1 | 2 | 3 | 4 | 5 | 104 | 1 | 2 | 3 | 4 | 5 |
| 5 | 1 | 2 | 3 | 4 | 5 | 105 | 1 | 2 | 3 | 4 | 5 |
| 6 | 1 | 2 | 3 | 4 | 5 | 106 | 1 | 2 | 3 | 4 | 5 |
| 7 | 1 | 2 | 3 | 4 | 5 | 107 | 1 | 2 | 3 | 4 | 5 |
| 8 | 1 | 2 | 3 | 4 | 5 | 108 | 1 | 2 | 3 | 4 | 5 |
| 9 | 1 | 2 | 3 | 4 | 5 | 109 | 1 | 2 | 3 | 4 | 5 |
| 10 | 1 | 2 | 3 | 4 | 5 | 110 | 1 | 2 | 3 | 4 | 5 |
| 11 | 1 | 2 | 3 | 4 | 5 | 111 | 1 | 2 | 3 | 4 | 5 |
| 12 | 1 | 2 | 3 | 4 | 5 | 112 | 1 | 2 | 3 | 4 | 5 |
| 13 | 1 | 2 | 3 | 4 | 5 | 113 | 1 | 2 | 3 | 4 | 5 |
| 14 | 1 | 2 | 3 | 4 | 5 | 114 | 1 | 2 | 3 | 4 | 5 |
| 15 | 1 | 2 | 3 | 4 | 5 | 115 | 1 | 2 | 3 | 4 | 5 |
| 16 | 1 | 2 | 3 | 4 | 5 | 116 | 1 | 2 | 3 | 4 | 5 |
| 17 | 1 | 2 | 3 | 4 | 5 | 117 | 1 | 2 | 3 | 4 | 5 |
| 18 | 1 | 2 | 3 | 4 | 5 | 118 | 1 | 2 | 3 | 4 | 5 |
| 19 | 1 | 2 | 3 | 4 | 5 | 119 | 1 | 2 | 3 | 4 | 5 |
| 20 | 1 | 2 | 3 | 4 | 5 | 120 | 1 | 2 | 3 | 4 | 5 |
| 21 | 1 | 2 | 3 | 4 | 5 | 121 | 1 | 2 | 3 | 4 | 5 |
| 22 | 1 | 2 | 3 | 4 | 5 | 122 | 1 | 2 | 3 | 4 | 5 |
| 23 | 1 | 2 | 3 | 4 | 5 | 123 | 1 | 2 | 3 | 4 | 5 |
| 24 | 1 | 2 | 3 | 4 | 5 | 124 | 1 | 2 | 3 | 4 | 5 |
| 25 | 1 | 2 | 3 | 4 | 5 | 125 | 1 | 2 | 3 | 4 | 5 |
| 26 | 1 | 2 | 3 | 4 | 5 | 126 | 1 | 2 | 3 | 4 | 5 |
| 27 | 1 | 2 | 3 | 4 | 5 | 127 | 1 | 2 | 3 | 4 | 5 |
| 28 | 1 | 2 | 3 | 4 | 5 | 128 | 1 | 2 | 3 | 4 | 5 |
| 29 | 1 | 2 | 3 | 4 | 5 | 129 | 1 | 2 | 3 | 4 | 5 |
| 30 | 1 | 2 | 3 | 4 | 5 | 130 | 1 | 2 | 3 | 4 | 5 |
| 31 | 1 | 2 | 3 | 4 | 5 | 131 | 1 | 2 | 3 | 4 | 5 |
| 32 | 1 | 2 | 3 | 4 | 5 | 132 | 1 | 2 | 3 | 4 | 5 |
| 33 | 1 | 2 | 3 | 4 | 5 | 133 | 1 | 2 | 3 | 4 | 5 |
| 34 | 1 | 2 | 3 | 4 | 5 | 134 | 1 | 2 | 3 | 4 | 5 |
| 35 | 1 | 2 | 3 | 4 | 5 | 135 | 1 | 2 | 3 | 4 | 5 |
| 36 | 1 | 2 | 3 | 4 | 5 | 136 | 1 | 2 | 3 | 4 | 5 |
| 37 | 1 | 2 | 3 | 4 | 5 | 137 | 1 | 2 | 3 | 4 | 5 |
| 38 | 1 | 2 | 3 | 4 | 5 | 138 | 1 | 2 | 3 | 4 | 5 |
| 39 | 1 | 2 | 3 | 4 | 5 | 139 | 1 | 2 | 3 | 4 | 5 |
| 40 | 1 | 2 | 3 | 4 | 5 | 140 | 1 | 2 | 3 | 4 | 5 |
| 41 | 1 | 2 | 3 | 4 | 5 | 141 | 1 | 2 | 3 | 4 | 5 |
| 42 | 1 | 2 | 3 | 4 | 5 | 142 | 1 | 2 | 3 | 4 | 5 |
| 43 | 1 | 2 | 3 | 4 | 5 | 143 | 1 | 2 | 3 | 4 | 5 |
| 44 | 1 | 2 | 3 | 4 | 5 | 144 | 1 | 2 | 3 | 4 | 5 |
| 45 | 1 | 2 | 3 | 4 | 5 | 145 | 1 | 2 | 3 | 4 | 5 |
| 46 | 1 | 2 | 3 | 4 | 5 | 146 | 1 | 2 | 3 | 4 | 5 |
| 47 | 1 | 2 | 3 | 4 | 5 | 147 | 1 | 2 | 3 | 4 | 5 |
| 48 | 1 | 2 | 3 | 4 | 5 | 148 | 1 | 2 | 3 | 4 | 5 |
| 49 | 1 | 2 | 3 | 4 | 5 | 149 | 1 | 2 | 3 | 4 | 5 |
| 50 | 1 | 2 | 3 | 4 | 5 | 150 | 1 | 2 | 3 | 4 | 5 |
Appendix B

Basic Tables
Appendix B

Table I

The Means, Standard Deviations, Standard Error of the Means, and the Coefficients of Relative Variability for all Groups on the Immediate Recall Test

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>SEM</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>17</td>
<td>35.90</td>
<td>5.34</td>
<td>1.33</td>
<td>14.87</td>
</tr>
<tr>
<td>B</td>
<td>17</td>
<td>35.82</td>
<td>3.75</td>
<td>.94</td>
<td>10.47</td>
</tr>
<tr>
<td>C</td>
<td>12</td>
<td>35.08</td>
<td>5.87</td>
<td>1.77</td>
<td>16.73</td>
</tr>
<tr>
<td>D</td>
<td>24</td>
<td>34.92</td>
<td>4.95</td>
<td>1.03</td>
<td>14.18</td>
</tr>
<tr>
<td>E</td>
<td>19</td>
<td>34.42</td>
<td>8.22</td>
<td>1.94</td>
<td>23.88</td>
</tr>
</tbody>
</table>

(Control)
Appendix B

Table II

The Means, Standard Deviations, the Standard Error of the Means, and the Coefficients of Relative Variability for All Groups on Delayed Recall Tests, with the Correlations Between Immediate Test and Delayed Test, the Differences Between the Means, the Standard Error of the Differences Between the Means, and the Critical Ratios.

<table>
<thead>
<tr>
<th>No of Cases</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval</td>
<td>1 week</td>
<td>4 weeks</td>
<td>8 weeks</td>
<td>12 weeks</td>
<td>12 weeks</td>
</tr>
<tr>
<td>Mean</td>
<td>34.29</td>
<td>31.71</td>
<td>30.58</td>
<td>30.75</td>
<td>31.21</td>
</tr>
<tr>
<td>S.D.</td>
<td>4.17</td>
<td>4.54</td>
<td>4.19</td>
<td>2.63</td>
<td>4.83</td>
</tr>
<tr>
<td>S.E.M.</td>
<td>1.04</td>
<td>1.14</td>
<td>1.26</td>
<td>.55</td>
<td>1.14</td>
</tr>
<tr>
<td>r</td>
<td>.66</td>
<td>.09</td>
<td>.41</td>
<td>.37</td>
<td>.43</td>
</tr>
<tr>
<td>Mean Diff.</td>
<td>1.61</td>
<td>4.11</td>
<td>4.50</td>
<td>4.17</td>
<td>3.21</td>
</tr>
<tr>
<td>S.E. Diff.</td>
<td>1.00</td>
<td>.31</td>
<td>1.70</td>
<td>.97</td>
<td>1.47</td>
</tr>
<tr>
<td>D/S.E. Diff.</td>
<td>1.6</td>
<td>13.0</td>
<td>2.7</td>
<td>4.3</td>
<td>2.2</td>
</tr>
<tr>
<td>Chances in 100 of a true diff.</td>
<td>94.</td>
<td>100.</td>
<td>99.7</td>
<td>100.</td>
<td>98.6</td>
</tr>
</tbody>
</table>