THE PSYCHOLOGICAL BULLETIN

LOGICAL LEARNING AND RETENTION: A GENERAL REVIEW OF EXPERIMENTS WITH MEANINGFUL VERBAL MATERIALS

BY E. L. WELBORN

Indiana State Teachers College

AND

HORACE ENGLISH

The Ohio State University

I. INTRODUCTION

Throughout the brief but extensive history of memory investigation, we find pleas for more attention to memory as it occurs in actual life. (For an example, cf. Meumann, 55). Of late there is increasing evidence that this appeal is being answered. One result is an emphasis upon "logical" memory or "memory for substance" in contrast with the serial learning of nonsense syllables. It now seems evident that any theory of learning is unlikely to prove acceptable unless it is based on investigations dealing with meaningful material. These investigations, however, have not been systematically reviewed as such since Whipple's material (75) was published in 1915. To make such a review is the purpose of this paper. We have selected from the considerable mass of published material those studies which help to bring to light the nature of "logical" memory.

By "logical" memory we mean memory for connected material measured in terms of the "substance" remembered. Logical memory is generally regarded as one end of a series at the other end of which is "rote" memory for purely nonsensical, entirely disconnected material. In this review we shall include some references to studies which contrast memory for highly meaningful material with rote memory even though they do not all measure memory in terms of

substance alone. We have, however, resisted the temptation to include studies of non-verbal learning, for until more secure interpretative formulae have emerged from the facts it is difficult to integrate data secured by too widely differing techniques.

II. SIGNIFICANT DIFFERENCES BETWEEN LOGICAL AND ROTE MEMORY

Ease in Learning. It has probably always been recognized that connected material may be learned more quickly than unconnected material of the same general character, though the fact has influenced learning theory strangely little. Earlier experimenters, such as Binet and Henri (11), obtained data showing greater ease in the learning of connected material, but most of these experimenters used a strictly verbatim criterion of learning. It is not easy to make a direct comparison between verbatim reproduction and logical memory in which only "ideas" are recalled. In logical memory percentages of correct reproduction obviously depend upon the scoring system and, if the recognition method is used, upon the degree of divergence between the original and the test items. To this must be added differences in the material learned, both as to content and form. The last were overcome by Jones and English (42) by measuring both logical and verbatim memory with the same subjects and the same passage. A test for "ideas" retained after one reading was followed by learning to the point of verbatim mastery. Ninety-one words of simple narrative were treated as yielding 31 "ideas." After one reading from 35% to 93% of these ideas were reproduced (with a mean of 76%). To master the ninety-one-word narrative verbatim required from 2 to 9 additional repetitions (mean 5.3). It may be added that other subjects, not reported above, failed entirely to master the verbatim task within appointed time and coöperative-tolerance limits, yet made nearly average scores in logical memory. It seems fairly clear, then, that just as verbatim learning of unconnected but meaningful material is easier than the learning of nonsense material and in turn as connected material is easier to learn verbatim than unconnected material, so the "ideas" are easier to learn than the exact words.

Retention. Comparison between logical and rote memory is more directly feasible with respect to retention since we can use immediate memory as a base for computing loss (or gain) at various intervals thereafter. Numerous studies furnish evidence contributing to such a comparison. We shall consider first a group of studies dealing

solely with logical memory. Despite variations in material and procedure, these studies show an unmistakable trend in the direction of better retention than the classical curve of Ebbinghaus shows.

Whipple (75) summarized earlier data showing that retention of connected material is better than that of discrete impressions. Henderson (37), for example, found that only 8% to 15% of his material was forgotten in 4 weeks. Yoakam (83) had pupils in grades 4 to 8 read a passage once and take tests of multiple choice and completion types. The mean score for all groups was 38.9% on immediate recall and 33.9% after 20 days, a loss of only 12.8%. Peterson (58) had 56 college students study a 250-word passage for 2½ minutes. By the reproduction method and scoring on "ideas," there was a loss after 1 week of only 15% of the immediate score. Greene (34) made a careful investigation using a large group of college students as subjects. With 1 presentation, the immediate retention was 53.1%, while that 1 week later was 41.2%, a loss of only 22.4%. Bassett (9) in a school testing program where the criterion was largely free from verbatim requirements obtained percentages of retention ranging from 86% at 4 months to 72% at 16 months.

The study of Dietze and Jones (19) is noteworthy because of its careful technique. Using prose passages over 1,000 words in length,

TABLE I
RETENTION SCORES OBTAINED BY THREE MEMORY STUDIES

Interval	Dietze & Jones (Prose)	Ebbinghaus (Syll.)	Radosawljewitsch (Poetry)
Immediate	56.7%	58.0%	96.0%
1 day	48.0%	34.0%	79.0%
14 days	36.3%	22.0%	30.0%
30 days	32.4%	21.0%	24.0%
100 days	30.0%		

presented once, large groups of school pupils as subjects, and highly reliable multiple choice tests, they obtained the data shown in Table I, together with other data for comparison.

The losses in terms of percentage of immediate retention between the immediate testing and that of the thirtieth day in these 3 studies are respectively 43%, 63.8%, and 75%. It thus appears that material read once and scored for substance was retained better after 30 days than other materials which had been repeated many times. Dietze and Jones, moreover, have devised a method of treating their data whereby all extraneous factors are supposed to be eliminated and the amount of actual forgetting determined. According to this method the retention at the indicated intervals was actually 90%, 76.2%, 59.3%, 51.7%, and 47.6%. It is of course not legitimate to compare these calculated percentages with those of other studies because none employs a similar tech-

nique, but in themselves they suggest that loss of substance is very gradual—certainly very much more gradual than the loss in case of rote memory. Indeed Trow (72) obtained a rise in retention score during 15 weeks, but he used only 2 subjects.

While the studies dealing solely with logical memory show in a general way that retention is much greater than in the case of rote memory, a number of other experiments have been carried out in such a way as to include both logical and verbatim memory in the same experiment. These supply material for more careful and exact comparison.

Sharp (64) and Henderson (37) found that words are forgotten to a greater extent than ideas. Greene (34) found that at the end of a week test items requiring an inference showed somewhat better retention than those requiring factual memory. Tyler (73) in a very careful study, obtained analogous results. Ability to answer questions demanding inference showed a much smaller loss over long periods than ability to recall or recognize factual material. Indeed in the former there was frequently a definite gain. Grant (33) showed that connected material is forgotten to a less degree than unconnected material. McGeoch and McKinney (53, 54) found incidental evidence of a difference between substance and verbatim memory in the 2 studies of retroactive inhibition in poetry and prose. Verbatim reproduction of poetry showed losses in 7 days greater in some cases than 50%, but retention in terms of combined verbatim and substance reproduction, a definite criterion for substance reproduction being used, was about 16% greater at the end of 7 days than immediately after learning. The loss in prose substance was only approximately 8% in the same interval.

English, Welborn, and Killian (21) made extensive comparisons of verbatim and substance memory for prose passages over 1,500 words in length. The procedure was unique in that comparison was made, not of 2 learning tasks as in other studies, but of the relative effectiveness of verbatim and substance memory in a single learning task. In the principal series of experiments the subjects read a passage and took a test of the true-false recognition type at various intervals. Two kinds of items were included in a single test: (1) verbatim items, which were practically perfect reproductions of statements from the passage, employing the vocabulary of the passage throughout; and (2) summaries or topic sentences employing a different vocabulary and covering sections of varying length in the passage itself. Separate scores were obtained for the 2 kinds of items. At no time were the subjects aware of the real purpose of the experiments or of the existence of 2 kinds of items. Throughout it was shown that during an interval of 4 to 14 weeks there was a consistent difference in median retention of the 2 kinds of items.

While the verbatim items showed moderate loss, the summary items showed either no loss or an actual gain. There were 2 independent investigations using different passages and tests, but the differences were essentially the same in all cases.

In a follow-up study English (22) computed not the median scores but the percentage of subjects whose scores indicated improvement in later trials over the immediate trial. Statistically reliable differences between the summary tests and the verbatim tests were again found.

The trend of these investigations clearly is that the curve for substance memory differs radically from that for rote memory. A few investigators contend, on the other hand, that the curve for substance memory is essentially similar to the Ebbinghaus curve.

Austin (6), who made a very extensive study, reported rapid loss in substance memory at first, but little loss between 2 and 4 weeks. She considered her results to be similar to those of Ebbinghaus. Jones (41) reported immediate retention, 62%, and fourteen-day retention, 30.6% as compared with 58% and 22% in Ebbinghaus' experiment. Davis and Moore (15) reviewed 61 memory studies and reached the conclusion that the curve of retention has the same general character in the case of nonsense and meaningful material. Unfortunately the studies are not listed so that we can determine whether the studies of meaningful material accepted "ideas" or demanded verbatim reproduction.

The number of important studies reporting data comparable with those of nonsense syllables is in fact still too small for any safe conclusion, especially in view of the great variation in procedure. Indeed, as Carr (13) has pointed out, there are as many curves as there are learning variables. There is none the less discernible a strong trend toward the conclusion that substance memory retention has a radically different dynamics from that of rote memory.

Correlational Data. Studies designed to throw light on the question of a general memory factor may first be reviewed for the differences which they reveal or do not reveal between logical and rote memory.

Woodrow (82) obtained intercorrelations of 6 memory tests, one of which was a test of substance memory, and found that the mean coefficient was .38. Johannsen, Stirling, and Levine (40) correlated scores in like and unlike pairs of school subjects and found some basis for stating that memory ability is specific for the type of material learned. Garrett (26) administered 8 memory tests to 158 students in Columbia College. The intercorrelations were very low and suggested that the tests dealt with relatively specific abilities. By the Spearman tetrad-difference technique he found a small but unreliable general memory factor. Anastasi (3) with a comparatively narrow range of tests

obtained evidence of a general memory factor. The same experimenter (4) later employed a wider range of tests for the same purpose. Her data are too extensive for brief summary, but we shall quote her conclusion: "The common factor previously found through a certain type of memory test cannot be regarded as a general memory factor extending through all forms of memory" (p. 45).

The second line of evidence deals directly with the degree of correlation between substance memory and various kinds of rote memory.

Wissler (81) correlated memory for a hundred-word passage with memory for digits and colors. For like pairs (*i.e.* prose with prose, or unconnected with unconnected) he obtained r 's of .29 and .39, while for unlike pairs the r 's were .03 to .05. Bennett (10), using various materials and only 9 subjects, obtained a mean coefficient between unconnected and connected material of $-.01$. Kitson (46) obtained a coefficient of .09 between logical material and numbers. Jones and English (42) obtained a coefficient of $.328 \pm .08$ between scores on ideas from a story read once and number of repetitions of the same story for verbatim learning. (In this experiment the attenuating effect of different materials upon the correlation was avoided; only the criterion of "memory" differed in such a way as to contrast substance and rote memory.) Anastasi's work (4) along this line is probably more thorough than any other. She obtained a coefficient of .647 (corrected for attenuation .74) between 2 tests of substance memory, but between substance memory and memory for syllables she obtained coefficients of .154 and .094. Welborn (74), whose study has already been described, obtained coefficients for pairs of similar tests ranging from $.664 \pm .04$ to $.884 \pm .02$, but for dissimilar pairs coefficients ranging from $.419 \pm .06$ to $.589 \pm .04$. Killian (44) obtained a coefficient of .39 in case of dissimilar pairs with a test having a reliability coefficient of .69. Smith and McDougall (67) contrasted the correlation of substance and nonsense syllable learning with the correlation of 2 forms of purely rote memory. In the 2 cases the coefficients were respectively .48 and .61.

The trend of all the correlational data may thus be fairly interpreted as demonstrating that different factors operate in substance and rote memory.

Other Differences. Fox (24) in reviewing certain experiments concluded that the enormous difference in the number of repetitions required to learn nonsense and meaningful material is in itself an indication of the disparity between the two. Killian (44) obtained some evidence, though inconclusive, that memory for his verbatim and summary items was differently affected by the number of repetitions.

Winch (79) studied the transfer of training in rote memory for groups of letters to substance memory for stories. His results are

inconclusive, but the study is mentioned as an example of the many possible ways of getting at the issue.

III. FORM OR ORGANIZATION AS THE BASIS OF SUBSTANCE LEARNING

The influence of the total meaning or pattern seems to have been noted first by Henderson. He concluded that the fact that all the topics are vaguely in the mind at once is "quite different from mere mechanical association."

Lewis (49) found that memory for hundred-word legends is largely a matter of the persistence of organized wholes. In his material he included irrelevant details to the extent of 37.5%; they accounted, however, for 63% of the omissions in recall. Laird, Remmers, and Peterson (47) found that in case of learning Anglo-Saxon words, additions, and dates, logical grouping gave much better results than learning in random order, that organized recall was better than unorganized, and that the balance in favor of organization became greater with the lapse of time. Key (43) also found that organization is an important factor. She showed that organization in memory includes two factors, viz., typification and unification. By the first she means that the essential details stand out prominently and that associations relating to the form as a whole tend to be retained. By the second she means that details are gathered into groups or units. Guilford (35) compared memory for number series that had plan or form with that for series without plan; also word lists that could be made into sentences with those that could not. His data showed a reliable difference in favor of material possessing form or pattern.

Aside from total pattern, detailed meanings, relations, inferences, etc., have been found to aid substance memory.

Foster (23) read a number of stories repeatedly to small children. By the use of the method of hesitating for the child to supply the next word, she found that correct response depended little on the familiarity of the words, but considerably on the closeness of relationship between the last word read and the one requested. Key (43) found that learning of meaningful material was 20 times as great as learning of non-meaningful material. Reed (61) studied the influence of associations in learning various types of material, and showed that easy prose such as *The Marble Statue* is learned far more quickly (111 sec. vs. 261 sec.) and retained much better (39.5 vs. 4.00 at the end of 2 weeks) than an abstract passage such as *The Origin of Ideas*. He concluded that the most valuable type of association is that based on meaning. Substance memory seems to be aided by reasoning ability. For example, Peterson (57) obtained a correlation coefficient of .40 between substance memory for passages and generalization and reasoning tests. Moreover, substance memory appears to be related to inventiveness. Winch (80) studied the effect of practice in memorizing short stories on ability to invent stories from a given list of words and found some transfer not due to common elements of material. The 2 abilities correlated from .276 to .749. Arnold (5) studied directly the factor

of association in memory for poetry and obtained evidence alleged to be incompatible with associationistic principles.

It has been found in a large number of studies that with the lapse of time various types of change occur in memory for substance.

Henderson (37), Austin (6), and Bartlett (8) noted that the material becomes more condensed and generalized with loss of original wording. Shaw (65), Binet and Henri (11), Henderson (37), and Lewis (49) found also that the accessory parts and accidentals are lost to a much greater degree than the fundamental ideas. But even more striking changes have been found. Binet and Henri (11) noted verbal assimilation, *i.e.* simplifying by substituting familiar for unfamiliar words. Construction or introduction has been found by Henderson (37), Key (43), Bartlett (8), Müller-Freienfels (56), and others. Henderson also found regrouping and simplification.

IV. RELATION OF SUBSTANCE MEMORY TO OTHER VARIABLES

Intelligence of the Subject. Several studies show that the more intelligent subject tends to do better than the less intelligent. Simpson (66) found that only 10% of a poor group surpassed the lowest 23% of a good group. English (20) with a somewhat similar procedure found that on prose passages, the superiority of a group of highly intelligent and socially advantaged English boys over an inferior group was "almost overwhelming." With verbatim recall of unconnected words the superiority of the advantaged group was less pronounced. Within each group, moreover, substance memory showed a considerably higher correlation with estimated intelligence.

Dietze (17) obtained mean coefficients between M.A. and memory score ranging from .55 for immediate tests to .25 for tests at 100 days. Wilson (77) compared retention of a story by dull nine-year-old, bright nine-year-old, dull twelve-year-old, and bright twelve-year-old children. The curves were practically parallel, the bright nine-year-olds and dull twelve-year-olds having practically identical curves in a middle position and the bright twelve-year-olds having one somewhat above and the dull nine-year-olds one somewhat below the middle curves. Wilson (78) obtained some evidence that brighter children do better with abstract and more difficult ideas than duller children. Grant (33) gave 10 memory tests, 3 of which were based on connected material. The coefficients between these three and intelligence were higher than those of any of the other tests. Bassett (9), in the case of retention of history, obtained a mean correlation with M.A. of .366. Greene (34) found that the upper quartile obtained their best scores after individual reading, while the lower quartile did best after lectures. With college students Garrett (26) obtained a coefficient of only .29 for logical memory *vs.* intelligence, but Foster (23) obtained a high correlation between M.A. and substance memory in young children. Good (32) found a greater difference between the upper and lower quartiles in performance requiring thought than in reproduction of ideas. Bolton (12) found a

close relation ($r=.695\pm.025$) between logical memory and intelligence while with the same 50 students the relation between "retentivity" and intelligence was only .065. Hegge (36), working with 98 subnormals at Vineland, obtained a correlation of .757 between logical memory and intelligence.

Rapidity of Learning. Lyon (51), Pyle (59), Henderson (37), Gates (29), Peterson (58), and Stump (70) found evidence that rapid learners retain connected material better than slow learners. Peterson found also that the advantage of the rapid learner increases in proportion to the difficulty of the material. The difference found by Stump is striking. Rapid learners lost little in 42 days, while slow learners lost as much as 60%. The correlation between rate of learning and amount retained was .71.

Scholarship. Wissler (81), Henderson (37), Pyle (59), Simpson (66), and Travis (71) obtained rather low correlations between substance memory and scholarship. King and Homan (45) obtained a number of coefficients with several groups and materials. The highest coefficient for any group was .38. They assert that logical memory is a "fair measure" of school standing. Pyle suggests that substance memory does not correlate very highly with scholarship because numerous other variables, especially study habits, influence scholarship.

Age. Henderson (37), Shaw (65), King and Homan (45), Pyle (60), and Dietze and Jones (19) present evidence that substance memory scores increase from the earliest ages tested to puberty.

Foster (23) obtained a correlation of .74 in case of a group of children having C.A.'s from 2-7 to 4-9, but for subjects in grades 7 to 12 Dietze obtained low r 's between C.A. and retention scores, the mean coefficients ranging from $-.02$ to $+.24$. Henderson and King and Homan found evidence that the increase continues through college. On the other hand Garrett (26) obtained "substantially 0" correlation in case of a group 16 to 26 years old, and Lodge and Jackson (50) found that college students under 25 did better than those over 25. Shaw and Henderson found that on the basis of later reproductions younger subjects retain as well as adults. King and Homan found that the per cent who do better on the second than the first reproduction increases from 19 in elementary school to 45 in college.

We may accept Dietze's conclusion that factual memory (by which he means substantially what we have called logical memory) shows decided increase from early life to puberty and that it increases more slowly from puberty to maturity.

Sex. Shaw (65) found girls superior to boys from the third grade to college; Pyle (60) found the same true from 8 to 15 years of age. On the other hand Foster (23) found boys consistently superior in case of her group ranging from 2-7 to 4-9 years of age, and Dietze (18) found that in a large

group in grades 7 to 12, when comparison was made by age groups, the boys did better. Pyle (60) found that the scores of boys increase up to 18 while those of girls increase up to only 15 years of age. Studies by Wissler (81), Gates (27), Lodge and Jackson (50), and Travis (71), agree on the point that women are superior to men in substance memory. In none of these studies, apparently, was the factor of physiological age held constant.

Material. Johannsen, Stirling, and Levine (40) correlated memory for literature, history, and science and concluded that memory ability is specific for the type of material. Reed (61) found that material with familiar associations has an advantage both for learning and retention. Binet and Henri (11) found that with increase in length of the material the number of words reproduced increases, but not in direct proportion. Also, there were more substitutions than omissions in the case of short sentences, but the reverse was true in the case of long sentences.

Serial Position. Dell (16) and Wilson (78) found that this factor is negligible in case of connected material. On the other hand Shaw (65) showed by various calculations that memory for the first part of a story, particularly the first sentence, is better than for the latter part. Also, Jersild (39) found that primacy has a high value in memory for a biography.

Mode of Presentation. Greene (34) found no significant difference in the case of college students between lectures and reading on immediate tests, but on retention tests at the end of one week, there was a highly reliable difference (99 chances out of 100) in favor of lectures. Kitson (46) obtained a coefficient of only .26 between logical material heard and seen. Russell (62) found that the advantage shifts from oral presentation in the fifth grade to silent reading in the ninth.

Feeling Tone and Temperament. A considerable number, perhaps a majority, of the studies of hedonic tone have reported substance memory, although seldom with much attempt to separate it from rote memory. These studies have been reviewed by Meltzer (54a) and need not be repeated here.

Attitude of the Learner. Fox (24) concluded from a review of experimental data that the desire or will to remember is the essential determining factor. Aall (1) tested memory for 2 stories and found a mean advantage of 13% for the intention-to-remember group. We have not had an opportunity to analyze Aall's data. Carr (13) expresses the conviction that in most memory experiments the amount recalled is a function of the expectation to be tested. Jones (41) compared retention of college lectures in case groups with and without tests immediately after learning. The effect of the tests was to keep the retention curve practically horizontal for 4 days; thereafter the decrement was slow during 8 weeks, at the end of which time those who took the immediate test did almost twice as

well as those who did not take the test. Bassett (9) obtained a mean correlation of .664 between interest in, and retention of, history. Fox (25) in a comparison of the whole and "mixed" methods of memorizing poetry found that the preference of the subject for certain of the poems had greater influence than the method used.

Guidance. Greene (34) compared guided reading with unguided reading and lectures by having the experimental group mark the test questions while they studied. On inference tests the retention score was 62%, while that for the control group was 10%. On information tests the per cent was 64 against 46. These differences were highly reliable.

Germane (30) found that reading preceded by a list of questions is better than rereading or reading followed by making a summary, and Yoakam (83) found that an initial test aids permanent retention. A test given before the completion of learning may be regarded as a form of guidance. Welborn (74) gave such tests and found a fairly reliable advantage in 7 comparisons out of 8. The familiar recitation method may possibly be interpreted as guidance. Gates (28) studied the effect of recitation on memory for biographies and found that for the optimal division of time (60% devoted to recitation) on immediate tests the difference in favor of the recitation method was 26.6% and on retention tests 3 or 4 hours later, 49.7%. Trow (72) found in a series of practice periods that presentation preceded by recall gave better results than several other procedures.

Number of Repetitions. While the evidence is neither extensive nor conclusive, it indicates that repetition is not of major importance in logical memory.

Killian (44) compared scores based on various numbers of readings from 1 to 4 and found small, but inconclusive, differences in favor of 3 readings. Good (31) found that 2 readings are not much more effective than 1 reading, and Henderson (37) found that a single attentive reading of a passage (135 to 180 words) is as effective as 3 minutes of study. Shaffer (63) and Jersild (38) have shown that an increase of scores occurs with repeated readings up to 4 and 5 respectively, but under conditions that raise doubt as to the true causes. Shaffer's subjects took a pre-test and a test following each reading. Jersild's results are difficult to evaluate because of the large number of variables in his experiment, but taken at face value, they do not show conclusive advantages for repetition. He concludes that "one of the least effective forms of emphasis is to repeat an item following its first presentation" and that "the benefit does not increase in proportion to the number of added repetitions" (p. 70). From the standpoint of study methods, Yoakam (83) and Dietze and Jones (19) proved that a single reading of a lesson is very inefficient.

In view of the Gestalt contention that frequency operates merely to make possible the acquisition of insights or meanings previously

not apprehended, it would be illuminating to have some of these studies analyzed item by item after each repetition. So far as we know, this has not been done with substance learning, although it is more appropriate for such a check than nonsense learning.

Spacing. The data for verbatim memory, which are fairly extensive, do not come within the limits of this review. Three careful studies of substance memory show an advantage for spaced learning.

Foster (23) compared 5 readings a day for 2 days with daily readings and found that the daily method gave better results especially in earlier reproduction. Austin (6) found that daily repetitions are better than 5 a day and that in a range from 5 per day to one every fifth day, daily and alternate daily repetitions are best, but contrary to Foster, she found that the advantage shows more on retention than immediate tests. Welborn (74) with 4 repetitions had spacings between them ranging from none (*i.e.* 4 consecutive repetitions) to 1 repetition every third day. He found that three-hour and daily spacing gave the best results.

*Reminiscence.*¹ Williams (76) and McGeoch (52) have made careful studies of this factor and have found evidence of its presence with tests combining verbatim and substance learning. Henderson (37), King and Homan (45), and Welborn and Killian (74) incidentally found evidence of reminiscence in substance learning. McGeoch (52) found that the mean scores of the entire group may conceal reminiscence and that the percentage of subjects exhibiting reminiscence and the amount shown by such subjects are adequate and valid measures. Following this cue, English (22) reworked some of Killian's data. Forty-six per cent of the subjects showed gains over the immediate scores on verbatim recognition after 24 hours, while 52% showed gains on summary recognition. Thereafter the percentage declined for verbatims but increased for summaries or remained practically constant. After 71 days in one group the percentage of subjects with higher scores than immediately after the presentation was 54 for summaries, as compared with 22 for verbatims.

Retroactive Inhibition. The effect of interpolated tasks on substance memory is largely unexplored. However, McGeoch and

¹ This term, introduced by Ballard (7), while justifiable on etymological grounds, cuts across the more common English usages. For want, apparently, of any better term, however, it is rapidly gaining ground for either (a) the recall or recognition of any item which has once failed of reproduction without overt intervenient presentation or practice, or (b) a net gain in recall or recognition for a number of items due to the mixture of recalled and of forgotten items under such conditions.

McKinney (53, 54) made 2 studies of this kind. In the case of poetry, when part of the same poem and also nonsense syllables were used as interpolated material, small amounts of retroactive inhibition were found in both cases after 15 minutes. The susceptibility to inhibition was considerably greater after 7 days than after 15 minutes and was enhanced when recall after 15 minutes was omitted. A similar study of substance memory of a prose passage showed approximately the same results. The authors believe that "these experiments increase the generality of the decrement phenomenon and give support to the theory of inhibition as a major condition of learning."

V. THE TECHNIQUE OF INVESTIGATION OF SUBSTANCE MEMORY

Control of Variables. We may well start with the assumption of Carr (13) that there are no constants and that variations in results are due to variations in conditions. The use of large groups of subjects is advantageous because the trend of results is not so much influenced by individual differences, daily fluctuations, etc., as it is in the case of small groups. The more important studies heretofore mentioned, including the earlier ones such as Henderson's, as well as many of the more recent ones, made use of large groups of subjects. The list of approximately 25 such studies will be omitted. In a few cases equated groups have been used. Greene's study (34) employed this procedure.

The problem of securing uniformity in difficulty of passages used in an experiment has been attacked in at least two ways aside from the use of the experimenter's judgment. Dietze and Jones (19) obtained the regression equations from data on their various pairs of passages and converted all scores into derived scores for one passage. Space does not permit the evaluation of their procedure. Jones and English (42) used the same passage for 2 kinds of testing, and English, Welborn, and Killian (21) based both kinds of test items used in their study on the same passage, thereby testing 2 kinds of memory for 1 learning task. The use of control groups for the purpose of eliminating the influence of prior learning is rather common, but the procedure of Dietze and Jones (19) for eliminating the effect of comprehension difficulties and poor habits of study, which has already been referred to, has been used by no one else so far as we know. The importance of the subject's knowledge as to whether and when he will be tested has already been mentioned; certainly this condition should be kept uniform in any

experiment. Practically all studies depend upon the subject's coöperation or motivation. The amount of exposure of the subject to the material has in some cases been regulated in terms of a time limit. Henderson (37) and several others used this method. The more common method and the one which seems to us to approach more nearly a true uniformity is that in terms of number of readings. Several of the most outstanding experiments, notably that of Anastasi (4), have employed only one reading.

Type and Length of Passage. There are enormous variations here, mostly unmeasured. The ease with which these factors are controlled—or seem to be—is, of course, a chief reason for their continued use. Travis (71) gives a useful methodological summary.

Test Technique. The method of reproduction has a certain logical justification in the case of substance memory and indeed it has been used in most of the pioneer studies as well as in some more recent ones. The usual scoring method is to divide up the passage into "ideas" and score one point for each "idea." There are two objections to this method. It precludes any investigation of reliability of measurement, and it is at the same time too subjective and not subjective enough. When the subject's reproduction varies in language or form from the original material, the court of last resort is the judgment of the experimenter. In some cases, *e.g.* Binet and Henri, substitutions have been rejected, but it is likely that in a majority of cases they have been accepted if approximately equivalent to the original. Henderson (37) found that the correlation between capacity for learning and for retention is shown better by scoring ideas than by scoring words. On the other hand Lodge and Jackson (50) and Travis (71) argue very strongly that scoring on "ideas" is not subjective enough. They contend that the method assigns too high a value to rote memory, that ideas should not be counted as equal in value, and that "judgment in organizing and condensing material, terseness and vigor of expression," etc., should be considered, whereas by the usual method the highest scores may be made by those who lack habits of selection and condensation. They prefer and use a qualitative scoring scheme based on a five-point ranking procedure as well as the usual scheme. Their data are too extensive for summary here. It seems to us that some combination of the 2 methods might be evolved through further experimentation.

Objective tests have been used very extensively in the more recent investigations, the commonest types being the recall or com-

pletion test and variations of recognition tests such as multiple choice and true-false. The first question to be considered is that of the reliability of measurements of substance memory by means of these tests. In recent years the practice of calculating reliability coefficients has become practically universal in all careful studies. We have at hand reliability data from a number of studies. An analysis of all of the data produced the following results.

(1) Reliability coefficients, 16 in number, obtained by correlation of chance halves or equivalent tests, range from .47 to .94, with a mean of .72. (In a number of cases the coefficient used is itself the mean of several in a whole study.)

(2) Repetition-of-test correlations were roughly grouped into short interval and long interval correlations. Those for short intervals, 8 in number, range from .47 to .94 with a mean of .79, while those for long intervals, 14 in number, range from .35 to .86, with a mean of .61.

The highest coefficients for multiple choice recognition so far reported seem to be those of Dietze and Jones (.90 to .94) and Anastasi (.861), and for recall that of Anastasi (.875). Outside of the field of substance memory, Strong (69) has shown that the reliability of recognition decreases rapidly with the lapse of time.

We can state, therefore, that in general the studies from which these correlations were obtained are sufficiently reliable, and we think that in the main the data for most of the important statements throughout the paper are reliable.

Two other important questions arise. Do recall and recognition tests measure the same functions? Is the recall type more difficult than the recognition type? Anastasi (4) has made one of the very few studies dealing with these issues in substance memory and probably the best study in the whole field. As to relative difficulty, her data show a negligible difference between the two types, and she says, "the alleged superiority of recognition over recall scores does not hold for logical memory . . . in which one is led to expect the greatest difference between the two processes according to previous writers" (p. 54).

As to whether recall and recognition are different processes, she obtained a coefficient of .23 in the case of 2 recognition tests with different material, but the coefficient for recall and recognition tests of the same material was .74. Also she reviewed other studies that attempt to establish differences between these 2 measuring procedures, particularly those of Achilles (2) and Lee (48), and found that their

data fail to show that recall and recognition are any more unrelated than recall tests for different kinds of material. Her final conclusion is that "material is more potent than method in determining inter-test correlations" (p. 54).

Unfortunately, control of the material used, as above noted, is peculiarly difficult in substance learning, though Bartlett (8) and Davis (14) contend that it is less so than with so-called nonsense materials. It has received so little attention that close comparisons between one study and another remain impossible. In regard to the number of repetitions, for example, there is for each type and length of material apparently a different optimum. Similarly for spacing, about all we can safely say is that for a given kind of material there is an optimal spacing; it is hazardous to set even approximate limits for this optimum. Indeed throughout the whole field one finds not only a lack of adequate data but even more a lack of incisive analysis. We are inclined to echo Snoddy's comment (68) that when to the simplest questions one receives a barrage of utterly contradictory answers, we can only conclude that we have not yet won through to a correct analysis and probably have not asked the right questions. Not a few have suggested that substance learning may have distinct laws from those governing nonsense syllables but only a few have devised experiments designed to discover those laws. The field is not new, but it is still wide open to fundamental research.

BIBLIOGRAPHY

1. AALL, A., Die Bedeutung der Zeitperspektive bei der Einprägung für die Dauer der Gedächtnisbilder. *Ber. über d. V. Kongress f. exper. Psychol.*, 1912, 237-241.
2. ACHILLES, E. M., Experimental Studies in Recall and Recognition. *Arch. of Psychol.*, 1917-1922, 6, No. 44. Pp. 80.
3. ANASTASI, A., A Group Factor in Immediate Memory. *Arch. of Psychol.*, 1930, 18, No. 120. Pp. 61.
4. ANASTASI, A., Further Studies on the Memory Factor. *Arch. of Psychol.*, 1932-1933, 22, No. 142. Pp. 60.
5. ARNOLD, F., The Initial Tendency in Ideal Revival. *Amer. J. Psychol.*, 1907, 18, 239-252.
6. AUSTIN, S. D. M., A Study of Logical Memory. *Amer. J. Psychol.*, 1921, 32, 370-403.
7. BALLARD, P. B., Obliviscence and Reminiscence. *Brit. J. Psychol., Monog. Suppl.*, 1913, 1, No. 2. Pp. viii+82.
8. BARTLETT, F. C., *Remembering*. London: Cambridge Univ. Press, 1932. Pp. x+317.
9. BASSETT, S. J., Factors Influencing Retention of History in Sixth, Seventh, and Eighth Grades. *J. Educ. Psychol.*, 1929, 20, 683-690.

10. BENNETT, F., Correlations Between Different Memories. *J. Exper. Psychol.*, 1916, 1, 404-418.
11. BINET, A., and HENRI, V., La mémoire des phrases. *L'Ann. psychol.*, 1894, 1, 24-59.
12. BOLTON, E. B., Relation of Memory to Intelligence. *J. Exper. Psychol.*, 1931, 14, 37-67.
13. CARR, H. A., The Quest for Constants. *Psychol. Rev.*, 1933, 40, 514-532.
14. DAVIS
15. DAVIS, R. A., and MOORE, C. C., Methods of Measuring Retention. *J. Gener. Psychol.*, 1935, 12, 144-155.
16. DELL, J. A., Some Observations on the Learning of Sensible Material. *J. Educ. Psychol.*, 1912, 3, 401-406.
17. DIETZE, A. G., The Relation of Several Factors to Factual Memory. *J. Appl. Psychol.*, 1931, 15, 563-574.
18. DIETZE, A. G., Some Sex Differences in Factual Memory. *Amer. J. Psychol.*, 1932, 44, 319-321.
19. DIETZE, A. G., and JONES, G. E., Factual Memory of Secondary School Pupils for a Short Article Which They Read a Single Time. *J. Educ. Psychol.*, 1931, 22, 586-598, 667-676.
20. ENGLISH, H. B., An Experimental Study of Mental Capacities of School Children, Correlated with Social Status. *Psychol. Monog.*, 1917, 23, No. 3, 266-331.
21. ENGLISH, H. B., WELBORN, E. L., and KILLIAN, C. D., Studies in Substance Memorization. *J. Gener. Psychol.*, 1934, 11, 233-260.
22. ENGLISH, H. B., Further Data on Reminiscence. *Psychol. Bull.*, 1935, 32, 688.
23. FOSTER, J. C., Verbal Memory in Preschool Children. *J. Genet. Psychol.*, 1928, 35, 26-44.
24. FOX, C., *Educational Psychology*. Chap. 6. New York: Harcourt Brace & Co., 1925.
25. FOX, C., Influence of Subjective Preference on Memory. *Brit. J. Psychol.*, 1923, 13, 398-404.
26. GARRETT, H. E., Relations of Tests of Memory and Learning to Each Other and to General Intelligence in a Highly Selected Adult Group. *J. Educ. Psychol.*, 1928, 19, 601-613.
27. GATES, A. I., Experiment on the Relative Efficiency of Men and Women in Memory and Reasoning. *Psychol. Rev.*, 1917, 24, 139-146.
28. GATES, A. I., Recitation as a Factor in Memorizing. *Sch. & Soc.*, 1917, 6, 743-749.
29. GATES, A. I., Correlation of Immediate and Delayed Recall. *J. Educ. Psychol.*, 1918, 9, 489-496.
30. GERMANE, C. E., The Value of the Controlled Summary as a Method of Studying. *Sch. & Soc.*, 1920, 12, 591.
31. GOOD, C. V., The Effect of a Single Reading *vs.* Two Readings of a Given Body of Material. *J. Educ. Method*, 1926, 5, 325-329.
32. GOOD, C. V., The Effect of Extensive and Intensive Reading on Reproduction of Ideas or Thought Units. *J. Educ. Psychol.*, 1927, 18, 477-485.
33. GRANT, M. E., Some Theories and Experiments in the Field of Memory. *J. Educ. Psychol.*, 1932, 23, 511-527.

34. GREENE, E. B., Certain Aspects of Lecture, Reading and Guided Reading. *Sch. & Soc.*, 1934, 39, 619-624.
35. GUILFORD, J. P., The Rôle of Form in Learning. *J. Exper. Psychol.*, 1927, 10, 415-423.
36. HEGGE, T. G., An Experiment with the Logical Memory of Subnormals. *Training School Bull.* (Vineland), 1929, 26, 82-86.
37. HENDERSON, E. N., A Study of Memory for Connected Trains of Thought. *Psychol. Monog.*, 1903-1904, 5, No. 6. Pp. iv+94.
38. JERSILD, A., Modes of Emphasis in Public Speaking. *J. Appl. Psychol.*, 1928, 12, 611-620.
39. JERSILD, A., Primacy, Recency, Frequency, and Vividness. *J. Exper. Psychol.*, 1929, 12, 58-70.
40. JOHANNSEN, D., STIRLING, M., and LEVINE, J., An Experiment on Types of Memory Ability. *J. Educ. Psychol.*, 1932, 23, 28-40.
41. JONES, H. E., Experimental Studies of College Teaching. *Arch. of Psychol.*, 1923-1924, 10, No. 68. Pp. 71.
42. JONES, M. G., and ENGLISH, H. B., Notional vs. Rote Memory. *Amer. J. Psychol.*, 1926, 37, 602-603.
43. KEY, C. B., Recall as a Function of Perceived Relations. *Arch. of Psychol.*, 1925-1926, 13, No. 83. Pp. 106.
44. KILLIAN, C. D., ENGLISH, H. B., and WELBORN, E. L., Studies in Substance Memorization. *J. Gener. Psychol.*, 1934, 11, 233-260.
45. KING, I., and HOMAN, T. B., Logical Memory and School Grades. *J. Educ. Psychol.*, 1918, 9, 262-269.
46. KITSON, H. D., The Scientific Study of the College Student. *Psychol. Monog.*, 1917, 23, No. 1. Pp. 81.
47. LAIRD, D. A., REMMERS, H., and PETERSON, L. J., Experimental Study of the Influence of Organization of Material on Retention. *J. Exper. Psychol.*, 1923, 6, 69-82.
48. LEE, A. L., An Experimental Study of Retention and Its Relation to Intelligence. *Psychol. Monog.*, 1925, 34, No. 4. Pp. 43.
49. LEWIS, F. H., Note on the Doctrine of Memory Traces. *Psychol. Rev.*, 1933, 40, 90-97.
50. LODGE, R. C., and JACKSON, J. L., Reproduction of Prose Passages. *Psychol. Clin.*, 1916-1917, 10, 128-148.
51. LYON, D. O., Relation of Length of Material to Time to Learn and Optimal Distribution of Time. *J. Educ. Psychol.*, 1914, 5, 1-9, 85-91, 155-163.
52. MCGEOCH, G. O., The Conditions of Reminiscence. *Amer. J. Psychol.*, 1935, 47, 65-89.
53. MCGEOCH, J. A., and MCKINNEY, F., Retroactive Inhibition in the Learning of Poetry. *Amer. J. Psychol.*, 1934, 46, 19-33.
54. MCGEOCH, J. A., and MCKINNEY, F., The Susceptibility of Prose to Retroactive Inhibition. *Amer. J. Psychol.*, 1934, 46, 429-436.
- 54a. MELTZER, H., The Present Status of Experimental Studies on the Relation of Feeling to Memory. *Psychol. Bull.*, 1930, 37, 124-139.
55. MEUMANN, E., *The Psychology of Learning. An Experimental Investigation of the Economy and Technique of Memory.* (Trans. by D. Baird.) New York: D. Appleton and Co., 1913.

56. MÜLLER-FREIENFELS, R., Studien zur Lehre vom Gedächtnis. *Arch. f. d. ges. Psychol.*, 1915, 24, 65-105.
57. PETERSON, H. A., Correlations of Certain Mental Traits in Normal School Students. *Psychol. Rev.*, 1908, 15, 323-338.
58. PETERSON, H. A., Class Experiment on Individual Differences in Memory. *J. Educ. Psychol.*, 1925, 16, 247-252.
59. PYLE, W. H., Retention as Related to Repetition. *J. Educ. Psychol.*, 1911, 2, 311-321.
60. PYLE, W. H., *The Examination of School Children*. New York: The Macmillan Co., 1913.
61. REED, H. B., Repetition and Association in Learning. *Ped. Sem.*, 1924, 31, 147-155.
62. RUSSELL, R. D., A Comparison of Two Methods of Learning. *J. Educ. Res.*, 1928, 18, 235-236.
63. SHAFFER, L. F., A Learning Experiment in Social Studies. *J. Educ. Psychol.*, 1927, 18, 577-591.
64. SHARP, S. E., Individual Psychology. *Amer. J. Psychol.*, 1899, 10, 329-391.
65. SHAW, J. C., A Test of Memory in School Children. *Ped. Sem.*, 1896, 4, 61-78.
66. SIMPSON, B. R., Correlation of Mental Abilities. *T. C. Contr. to Educ.*, 1912, No. 53.
67. SMITH, M., and McDUGALL, W., Some Experiments in Learning and Retention. *Brit. J. Psychol.*, 1919-1920, 10, 199-209.
68. SNODDY, G. S., *Evidence for Two Opposed Prophecies in Mental Growth*. Lancaster, Pa.: Science Press Printing Co., 1935.
69. STRONG, E. K., JR., The Effect of Time Interval upon Recognition Memory. *Psychol. Rev.*, 1913, 20, 339-372.
70. STUMP, N. F., A Classroom Experiment in Logical Learning. *J. Appl. Psychol.*, 1927, 11, 117-126.
71. TRAVIS, A., Reproduction of Short Prose Passages: A Study of Two Binet Tests. *Psychol. Clin.*, 1915-1916, 9, 189-209.
72. TROW, W. C., Recall vs. Repetition in Learning of Rote and Meaningful Material. *Amer. J. Psychol.*, 1928, 40, 112-116.
73. TYLER, R. W., What High School Pupils Forget. *Educ. Res. Bull.*, 1930, 9, 490-492.
74. WELBORN, E. L., ENGLISH, H. B., and KILLIAN, C. D., Studies in Substance Memorization. *J. Gener. Psychol.*, 1934, 11, 233-260.
75. WHIPPLE, G. M., *Manual of Mental and Physical Tests*. (Sec. Ed.) Baltimore: Warwick & York, 1915. Part 2, pp. 205-223.
76. WILLIAMS, O., A Study of the Phenomenon of Reminiscence. *J. Exper. Psychol.*, 1926, 9, 368-387.
77. WILSON, F. T., Learning of Bright and Dull Children. *T. C. Contr. to Educ.*, 1928, No. 292.
78. WILSON, F. T., A Comparison of Difficulty and Improvement in the Learning of Bright and Dull Children in Reproducing a Descriptive Selection. *Genet. Psychol. Monog.*, 1931, 9, 395-435.

79. WINCH, W. H., Transfer of Improvement in Memory in School Children. *Brit. J. Psychol.*, 1909-1910, 3, 386-405.
80. WINCH, W. H., Some Relations Between Substance Memory and Productive Imagination in School Children. *Brit. J. Psychol.*, 1910-1911, 4, 95-125.
81. WISSLER, C., Correlation of Mental and Physical Tests. *Psychol. Monog.*, 1899-1901, 3, No. 6. Pp. 62.
82. WOODROW, H., Effect of Type of Training on Transference. *J. Educ. Psychol.*, 1927, 18, 159-172.
83. YOAKAM, G. A., The Effect of a Single Reading. *Nat. Soc. for Stud. of Educ. Yearbook*, 1921, 20, Part 2, Chap. 6.