INTERRELATIONS BETWEEN LEARNING ABILITIES IN VARIOUS SITUATIONS.

Carol Frances Lundie.

Thesis submitted in part fulfilment of
the requirements for the degree of
Master of Arts, in Psychology.

Rhodes University College,
Grahamstown,
January 1942.

A C K N O W L E D G M E N T S.

The writer wishes to thank Professor E.H. Wild, Head of the Psychology Department at Rhodes University College, for his patient assistance and constant guidance. While engaged in this research the writer held a Muirhead Bursary, for which she acknowledges her indebtedness.

She also expresses her appreciation to the seventy women students at Rhodes University College, who, by their willing co-operation to act as subjects, made the execution of this investigation possible.

N O T E.

The numbers given in brackets, after the names of the various authors, relate to the alphabetical list of references at the end of the study.

In this list of references, in the case of journals, the name of the journal is given first, then the date of the journal then the volume, page number, and finally the title of the study referred to.

CONTENTS.

I N T R O D U C T I O N.

It is a matter of considerable theoretical and practical importance to determine whether learning ability is general or specific. In everyday life one often hears remarks of the following nature: "My niece will pick that up very easily, she is such a quick learner", or "Poor Johnnie has failed his class this year, but then he always did lack learning ability". These remarks are occasionally qualified by such statements as : "She is such a quick learner where anything involving needlework is concerned", or "He always did lack learning ability in school subjects". These qualifications, although they may cover a large range of activities, as in the second example, imply that the individual referred to has a different degree of learning ability in another sphere.

Not only do the majority of doting aunts and disappointed parents speak of learning ability as if individuals differed consistently in amount thereof, irrespective of what is learned, and the circumstances under which the learning occurs; but the existing laws of learning, formulated by some eminent psychologists, tacitly assume that there is a learning process, not different learning processes for different tasks. These psychologists imply by their statements an underlying general learning ability. There is a corresponding assumption of a general learning ability behind the practice of

intelligence testing; Thorndike (58 p.258), however, has pointed out that it is not known to what extent any test of intelligence is a measure of the ability to learn the elementary school subjects, or to learn to add better, or to learn a code, or anything else; that is, it is not known exactly to what extent learning ability is general or specific.

An individual can learn a great variety of things at different times, and under different circumstances. What we would like to know is whether learning ability is general in the sense that individuals differ consistently in respect to rate of learning, irrespective of what is learned, and of the circumstances under which the process occurs. Can one speak of quick learners and slow learners, without the qualification of mentioning what is being quickly or slowly learned? Is learning ability general or specific to the particular task being learned?

The problem can be compared with that of intelligence. It used to be assumed that there was a general ability called intelligence, but much work has been done of recent years, by Spearman, Thomson, Thorndike, Thurstone and others, on this problem, so that with regard to intelligence, there are now three theories: (a) that intelligence is a general ability; (b) that abilities run in groups; and (c) that one's abilities are highly specific. Present opinion seems to favour the general and group factor theories. Educational psychologists seem to assume that learning ability is as general as intellectual ability was assumed to be, and speak of fast and slow learners, without mentioning what is being learned. (Husband 30.)

Interest in the nature of learning ability is now appearing to develop, and the work that has already been done on this problem does not support the assumption of a general learning ability. The investigators concerned, however, seem dissatisfied with their results. They write as if there should have been higher correlations between learning abilities than they found, and it was this dissatisfaction on their part which stimulated the writer to undertake this present research. The repetition of an earlier investigation has usually been considered beneath the dignity of a self-respecting Master of Arts candidate! In fact psychological research generally has over emphasized exploration and under exphasized verification. What is needed in psychology is not mere blind repetition, but repetition with insight: A return to the original meaning of the word <u>research</u>. We might well heed the appeal of Ross (49), who writes: "Psychology has had too much search, and not enough <u>re</u>-. Novelty may be noteworthy, but let us make repetition also respectable". If methods can be designed to overcome some of the defects in previous investigations, then this repetition will be valuable. In fact psychologists should give up their present tendency to accept the results of investigators before they have been corroborated by the results of repetition by other researchers.

Before giving a survey of the work done by previous investigators in the field of learning ability, the writer would like to quote Pyle(47 p.155),

to show that the theorists are wanting, and waiting for information on the relationships between various kinds of learning abilities. He asks the following challenging questions: "To what extent is learning capacity general, and to what extent is it specific? What evidence is there that there is a general factor in learning? What evidence is there that there is always a specific factor? How can the general factor be best measured?" Few people have followed up these questions and the evidence which has been forthcoming has been far from decisive, as will be seen in the following chapter. Hence this research is justified if it can throw further light on some of these questions, even if it only corroborates the results of the few previous investigators into this field.

CHAPTER I.

A SURVEY OF PREVIOUS STUDIES INTO THE INTERCORRELATIONS AMONG LEARNING ABILITIES.

Before attempting to draw up a test battery, it is necessary to make a critical survey of available literature bearing on the problem of interrelations between learning ability in various situations, so that one can profit by previous investigators' experience, and find out aspects of the problem which need special attention.

A. Early Investigations.

Hall (23) summarises the work of former investigators into the field of human learning, which yields some evidence which might suggest answers to the question of generality or specificity in learning ability. More particularly this summary bears on the question: are the correlations between different measures of improvability of high or low magnitude? Hall obtained correlation coefficients between scores obtained on various learning tasks from Brooks (4), Chapman (11), Garrett (17), Gates (18), Gunlach (22), Haught (24), Heron (25), Race (48) and Thorndike (57) the variety of tasks included colour-naming, cancellation, opposites, addition, mental multiplication, typewriting, digit-symbol substitution, Turkish-English vocabulary, code learning, rational learning, checker puzzle, stylus maze, inverted writing, number completion, tapping and word building. Five of the investigators measured learning in terms of the improvement or gain from the initial to the final trial,

while the remaining four used scores based on the total practice period. The eighty-four correlation coefficients found by these investigators, Hall arranges in tabular form, showing their distribution. They range from one which lies between ·70 to ·79, to one between - ·40 and -·49. The majority of coefficients lie between· 00 and ·39; the median lies between ·20 and ·29, and the median of the positive coefficients is+ ·25. Only one-third of the total number of coefficients is significantly different from zero (taking the difference between the coefficient and zero as four times its probable error), and about fourteen percent are greater than ·5. From these results it would seem that a general learning ability if such exists, is of slight importance in determining learning performance. However, it is necessary to differentiate between performance in the learning situation and the actual learning process. Learning ability has to be measured in terms of learning performance, and the score obtained on such a performance is doubtless contaminated by many factors extraneous to the learning ability or abilities of the subject. Therefore, to some extent, the lowness of the correlations may be due to the influence exerted upon the scores by these extraneous factors. Hall enumerates some of the irrelevant factors which may raise or lower inter correlations between learning tasks. As the part played by these factors on the size of the correlations, is vital to the problem, a summary of Hall's list follows.

B. Irrelevant Factors which may raise or lower inter-correlations between Learning Tasks.

1. Range of Talent. It is a well-known

statistical principle that the correlation coefficient is systematically lowered by restricting the range of talent sampled. If a representative sample of the total population were tested, instead of college students, who have been used as subjects in most of the previous investigations, higher correlations would probably be obtained. How much the r's would increase by doing this, is highly speculative.

2. Correlation due to irrelevant factors.

If a variable, such as age, sex, nationality, education or socio-economic status, be correlated with several learning performances, the intercorrelations between the learning scores will be spuriously increased.

3. Unreliability of Measurement.

The crude r's obtained can be corrected for attenuation, if the reliability coefficients of the learning tasks are known. In Hall's Table I, taking the crude r's for which "true" r's can be calculated, he finds an increase in the median value from .29 to .47. Hall concludes from this that the unreliability of measurement is an important cause of low inter-correlations. However the methods employed in calculating the reliability coefficients of measures of learning tasks can be severely criticised, which means that the r's may have been spuriously increased by correcting for attenuation in this way. For further discussion on this topic see page 19 .

4. Number of Subjects.

The stability of the obtained r's is in part a function of the number of subjects used. Correlations based

on fewer than thirty to fifty subjects are inconclusive, because of the large probable errors, and because the usual probable error formulae do not hold when applied to less than twenty-five cases.

5. Measure of Learning.

Difference in the magnitude of the r's reported in Hall's Table I, may be due in part to variation in the methods employed for measuring the results of practice. As has already been stated, five of the investigators measured learning in terms of improvement or gain from the initial to the final trial, the remaining four using scores based on the total practice period.

6. Number of trials.

Intercorrelations between scores on a battery of learning tests are affected by the number of trials on which the scores are based. According to Hollingworth (27), the correlation between two tasks worked on for five trials will be lower than the correlation for the same task worked on for ten trials.

7. Types of learning tasks used.

A critical test of the existence of a general learning ability would consist of an investigation employing a sample of learning tasks representative as to content and complexity. By judicious sampling of learning tasks with very similar content, it should be possible to obtain r's of a high magnitude, but one could not argue therefrom for a general learning ability. On the other hand one could make out a very good case for specialised learning abilities if low correlations were obtained between very similar learning tasks.

8. Previous Practice.

The ability evidenced by an individual in a learning task will depend on the amount of previous practice he has had on the task. If one individual has had one unit of previous practice on Task A, and two units on Task B, and a second individual has had two units of practice on A, and one unit on B, their relative initial position on the two tasks will be reversed. Should this condition of varying amounts of earlier training prevail for all subjects on the same task, and for the same individual on different tasks, the correlations between the several measures will be lowered. Numerous attempts have been made to devise a task which will eliminate the influence of previous training, either by starting all individuals at zero ignorance, or at some common point on the learning curve; but none have been really successful. Some tasks, such as mazes, seem to minimise the influence of past experience.

9. Positive and Negative Transfer of Training.

Positive transfer is an increase in efficiency in learning Task B as a result of having learned Task A previously. Negative transfer of training is defined as a decrease in efficiency in learning Task B, as a result of having learned Task A. Therefore having learned Task A may affect the results on Task B either favourably or unfavourably, and it is probable that subjects will be differentially affected by transfer, in which case the true relationships between the learning tasks will be masked.

10. Understanding of the directions.

11. Emotional adjustment to the learning situation.

12. Motivation.

13. Influence of other extraneous factors on the learning process.

A likely result of failure to control one or all of these factors is that they will have a differential effect on the same individual in different learning situations, consequently his performance will vary from task to task and the correlations between performances will be reduced.

Having elaborated these thirteen factors which influence the intercorrelations between learning performances, Hall says that an individual's learning score may be represented by the following equation: Score on learning task = f(learning ability, irrelevant factors, unreliability of measurement, positive or negative transfer, previous practice, type of material used, measures of learning employed, number of trials, understanding of directions, emotional adjustment, motivation, and daily variations.) Another important, and often neglected variable affecting the scores obtained on learning tasks in a differential manner, is variation in work methods. Attention has been drawn to this by Seashore (51) and van Dusen (16).

Of these fourteen variables, ten probably lower the correlation between learning tasks, one raises the correlation, and three may either raise or lower the intercorrelations. According to this analysis there is every reason to believe that the

inter-correlations would rise, should these irrelevant variables be eliminated.

In his investigation Hall states that some, but not all of these factors were controlled. A summary of his investigation follows.

C. Hall's Investigation.

The four learning tasks chosen were: A stylus maze (SM); the Petersen rational learning test (RL); a list of nonsense syllables (NS); and a punchboard maze (PB); chosen because of the frequency with which they had previously been used, known reliability, and apparent dissimilarity of content. As subjects a hundred first year college women, who volunteered to act as subjects, were used. They reported once a week on the same day and at the same hour, to practise one task each week; each task being practised for fourteen trials. Errors, or in the case of the nonsense syllables, the number right, for each trial were the only records kept. Explicit verbal directions, and thereafter a short fore-practice on each task was administered. No attempt was made at motivating the subjects, beyond asking them to do as well as they could. The measures of learning employed were total errors or number right for the fourteen trials, and improvability or learning ability was measured by the difference between the score on the first two trials and the last two. (1 + 2) - (13 + 14).

The reliability coefficients for the total error score were obtained by summing the errors made on the odd and even trials, and correlating these scores. The Brown-Spearman formula was applied to

give the reliability of the entire fourteen trials, the coefficients obtained in every case being above .9. The reliability coefficients for gains were obtained by correlating the gain from trial one to thirteen with the gain from trial 2 to 14, but the reliability coefficients with the exception of that found for nonsense syllables, were low in this case. For further discussion on reliability see page 19.

The intercorrelations for total error scores, and for improvement are as follows:

	Total errors.		Gains.
	Crude	Corrected for attenuation.	Crude.
PB vs.RL	.38	.40	-.21
PB vs.SM	.28	.30	.20
PB vs.NS	.33	.34	.03
RL vs.SM	.27	.29	-.02
RL vs.NS	.18	.18	-.12
SM vs.NS	.11	.11	.07

TABLE 1.HALL'S INTERCORRELATIONS FOR TOTAL ERRORS & GAINS.

None of the coefficients using gains as the measure of learning is significantly different from zero. The intercorrelations using the total error score as the criterion are all positive. Two are not significantly different from zero, and the remaining four do not indicate a very large community of function between the performances.

Hall attempts to show in how far the thirteen irrelevant factors have been controlled in his investigation. Unreliability of measurement, number of subjects, type of material used, measure of learning and number of trials, he maintains, were

so controlled as to constitute a fair test of the existence of a general learning ability. Correlation due to irrelevant factors, positive and negative transfer, understanding of directions, emotional adjustment and daily variations, were partially controlled. However, previous practice and motivation were definitely not controlled. In spite of the low correlation coefficients obtained, Hall believes it highly probable that a general learning ability as indicated by the presence of positive correlation between pure measures of learning, would appear, were one able to devise a set of tasks free from the differential effects of previous practice, and in the learning of which motivation could be made the same for all subjects. In conclusion Hall writes: "An original investigation in which some of the factors irrelevant to the learning process were eliminated, yields correlations of no greater magnitude than those obtained by the former investigators. Although the present and previous studies point to a high degree of specificity in learning, the writer believes that a general learning ability of some importance might be discovered were we able to control the differential influence of motivation and previous practice in the learning situation".

D. Husband's Investigation.

Husband (30), following up the work done by Hall, carried out an investigation involving many more tests than Hall's. The tests utilised were the following: (1) Learning names and faces; (2) Spool-packing; (3) memory for names and faces;(4) rational

learning (a mental maze test); (5) Mirror drawing; (6) Auditory prose; (7) cancellation; (8) maze learning; (9) Persian-English vocabulary, visual; (10) memory for prose passage; (11) visual prose; (12) code substitution; (13) pursuit rotor; (14) incidental memory for code; (15) card-sorting; (16) Hindu-English vocabulary, auditory; and (17) visual prose memory.

On most of the tests only five practice trials were given. One hundred college students were used as subjects, being individually tested in three sessions of an hour apiece, one week apart.

The correlations obtained from this investigation are shown in Table 2. As can be seen they are all extremely low. They range from -·16 to +·53, with a median coefficient of ·13. Two-thirds of the coefficients lie between -·20 and+·20, and are not significantly different from zero.

Test No.	2	3	4	5	7	8	9	10	12	13	14	15	16	17
3	·18													
4	·16	·04												
5	·04	·11	·01											
7	·09	·04	·10	·38										
8	·06	·00	·38	·25	-·11									
9	·17	·26	·23	·13	-·05	·32								
10	-·06	·37	·27	-·04	-·06	·08	·23							
12	·22	·06	·21	-·13	·07	-·05	·33	·09						
13	·16	·03	·03	·27	·13	·21	-·06	·10	-·06					
14	·03	·28	·05	·05	·06	·20	·16	·09	·04	·22				
15	·26	·25	·23	·24	-·04	-·04	·26	·29	·45	·22	·14			
16	·06	·25	·25	·00	-·04	·22	·53	·05	·17	·19	·20	·19		
17	·04	·16	·24	·23	-·03	·22	·30	·26	·11	·04	·18	·31	·27	
Intelligence	·10	-·01	·02	·22	-·13	·19	-·10	-·05	·19	·05	-·15	-·06	-·16	·29

TABLE 2. INTERCORRELATIONS AMONG LEARNING TASKS (AFTER HUSBAND). (the numbers indicating the tests refer to the numbers of the test quoted on pare13 & 14).

Husband sub-divides his intercorrelations into three groups, composed of tests of somewhat similar demands. The median among the motor tests is +·19, among the rote learning tests is +·20, and it rises to +·25 with the three correlations among ideational tests.

Examining Hall's list in terms of this investigation, Husband regards the following points as relevant: (a) Limited range of talent certainly applies to this investigation. College students of the same age, educational background etc., were used as subjects, which probably lowers the correlation coefficients obtained. (b) Similarity in age and culture possibly serves to raise the correlations. (c) Unreliability of measurement applies, as doing seventeen tests in three hours demanded shortening many to an undesirable degree. (d) The measure of learning used was the total scores in all trials, on the assumption that everyone started with an equal skill, and the faster learners had better scores as time went on, hence their totals should have been lower in time or greater in quantity than those who did not improve so much. Husband admits that this reasoning is theoretical rather than practical, and points out that some subjects caught on to the nature of the test earlier than others, and some were aided by more rapid eye or hand co-ordination, which permitted better scores due to native or previously acquired skill, and not necessarily through learning.

Husband concludes his article by saying: "The whole trend of evidence is such as to suggest very strongly that we must speak of learning abilities (plural) and not of learning ability (singular), as

if it were a general ability. Unless there is a great deal of overlapping in the nature of the tasks, interrelationships are very low".

Before this statement can be accepted further experimentation is necessary, and the theorists are not likely to give up speaking of "learning ability", until conclusive evidence has been put forward that a general learning ability does not exist. It must also be remembered that even with a relatively low range of correlations, general factors may present themselves if factorial analysis is applied to the table of intercorrelations.

E. Criticisms of these two investigations.

(a) Hall's Study.

(i) The Tests used. Hall states that one of the reasons for choosing the four learning tasks used in his investigation, was their apparent dissimilarity of content in order to sample as widely as possible the various so-called types of learning.

On looking at the tests it is found that they consist of a 40-T-unit stylus maze; a rational learning test requiring the subject to associate numbers with letters; a list of nonsense syllables to be memorised; and a punch-board maze into the face of which are bored thirty pairs of holes, the subject being required to go through the thirty pairs with a stylus, selecting the one of each pair which sounded a buzzer. These tasks, in the writer's opinion, seem to have a certain similarity of content. There are two mazes, neither of which allows of much spatial relation eduction. In the punch-board maze the

subject must learn the correct holes by repetition - i,e. the task is one involving mainly "facilitation" (See Chapter II, page 26). The same applies to a T-maze. The nonsense syllable test involves rote memory, a knowledge of the syllables being obtained by repetition - i.e. it is also a "facilitation" task. The so-called rational learning test probably gives more room for forming associations between the letters and the numbers, and more room for relation eduction; but subjects need not necessarily adopt such work methods, and the task can probably be done largely by repetition, especially as there are only sixteen pairs to be associated. Thus it would seem that the tasks are all of the kind that Husband would call "rote-memory" tests, and of the type which have been called "facilitation" tasks in this study. The two "motor" tasks seem to give little opportunity for motor learning of a manipulative kind, and none of the tests seem to give much opportunity for learning involving relation or correlate eduction. The nature of Hall's tasks probably accounts for the fact that his correlation coefficients are on the whole higher than those found by Husband.

(ii) Measures of learning. In computing the intercorrelations between the performances, Hall employed two measures of learning: Total errors or number right for the fourteen trials, and improvability as measured by the difference between the score on the first two trials and the last two (1 + 2) - (13 + 14). The former measure assumes that all subjects started with about an equal skill, and that the faster learners

had better scores as time went on, hence their totals should have been lower than for those who did not improve so much. From this investigator's experience she knows that subjects start with anything but equal skill (as represented by the score on the initial couple of trials). This measure of Hall's is therefore taking account of actual starting performance on the task, as well as of learning ability. (See Chapters II and VI.)

The second measure is a better one for guaging "learning ability", except that the first score should perhaps not have been taken into consideration, as Hall himself writes concerning the notorious unreliability of initial scores on learning tasks. Why not use (2+ 3) - (13+14)? This measure can be criticised on the grounds that it does not take into account the way in which the time reduction has taken place, whether regularly or irregularly, more quickly at the beginning, or at the end of fourteen trials etc. More discussion on this point will be found in Chapter VI.

It is from the correlations found by using the first measure that Hall draws most of his conclusions, and the size of these correlations are in part determined by the subjects "actual ability" on the task, as distinct from "learning ability". It would seem, from a comparison of the correlations obtained by using gains as the criterion of learning ability, with those obtained by using total time, that the latter measure tends to raise the correlation coefficients. As this measure does not take only learning ability into account, the correlations obtained

do not give a true reflection of the interrelationships of learning abilities, as they have been spuriously increased by taking "actual ability" into account, and its use can be criticised.

(iii) Method of measuring reliability and correcting for attenuation.

To find the reliability coefficients for the total error score, the errors made on the odd trials, and the errors made on the even trials were summed, and these scores correlated. This method of measuring reliability in the field of learning has been severely criticised by H.A.Carr (5). He points out that if the distribution of the scores for each of the trials is entirely a function of chance factors, the inter-trial correlation will approximate zero. If the distribution is wholly a function of the constitutional differences of the individuals; the correlation will be 1.00. He continues: "The logical validity of this method is based upon a tacit assumption i.e. that the size of the correlation is an exclusive function of these two opposite influences. In other words it is based upon the assumption that there is no third factor present which will account in part for the size of the correlation.

"What is a learning process? It is a process in which the performance of an individual on the tenth trial is a function of the preceding trial. It is a process in which the individual's performance changes from trial to trial in virtue of the preceding trials. It is a process in which the scores for the successive trials are so casually tied together that the scores

for the odd and even trials must necessarily exhibit some sort of correlation with each other. In all learning situations, the size of the obtained correlation must thus be a function, in part of what I shall hereafter call 'the learning function'."

By this method Hall finds high reliability coefficients, all being above .9. It is interesting to note that when reliability coefficients for gains were obtained, they proved to be low. Now this second method of finding the reliability of the tests, seems a more valid one to the writer. It is entirely arbitrary and a matter of expediency to the investigator, whether one takes the subjects' learning ability as improvement made from trials 1 to 13, or trials 2 to 14, or say 3 to 21 for that matter. If the test is reliable as a test of the learning process, the correlations between these methods should be fairly high; therefore the fact that Hall obtained low intercorrelations by this method, seems to indicate that the tests are not as reliable as one may be led to believe by using the odd and even trial method. A mistake Hall made, however, was in taking the initial score into consideration. He admits that the first trials of almost any learning task are notoriously unreliable. If he had taken as his criterion for learning ability (2 + 3) - (13 + 14), and correlated the gain from trial two to trial thirteen with the gain from trial three to fourteen, (2 - 13) vs. (3 - 14), he would probably have obtained higher reliability coefficients.

(b) Husband's Study.

(1) The Tests used. Husband uses more

tests than Hall, and the variety is also greater. Some of the tests are purposely made very similar, such as the "Persian-English, visual", and the "Hindu-English, auditory" vocabulary tests. The tests are divided, rather arbitrarily into motor, rote learning and ideational tests, but they do seem to cover more different types of learning than Hall's Tests do. Tests 10 and 17, the writer thinks, can be criticised as being of rather a different type from the other tests used. They test the subject's memory for a piece of prose, and do not lend themselves for measurement in terms of improvability; therefore she does not think they should be classed as "learning" tasks, but rather as "memory" tests.

It is also doubtful whether a cancellation test tests learning ability, or merely the powers of attention of the subject. If much improvement takes place, it probably points to the fact that the subject was not attending closely enough during the first trials, and is improving in this respect. (Whipple 60).

(ii) Number of trials. Doing seventeen tests in three hours meant shortening the tests to an undesirable degree. On many of the tests the subject was only given five trials, and as Hall points out from Hollingworth (27), the correlation between two tasks worked on for five trials will probably be lower than the correlation for the same tasks worked on for ten periods. (For further discussion on this point, see Chapter VIII, pg. 145).

(iii) Measure of Learning. The measure of learning used by Husband was also the total scores of

the subjects in all trials, which has already been criticised.

F. Summary.

Earlier studies of the relationships between measures of learning give surprisingly low correlations, from which the conclusion has been drawn that performance in the learning situation seems to be very largely specific to the type of task employed; but later investigators felt that the subjects' performance had in part been determined by factors other than their learning ability. In spite of improved methods of technique, and attempts to control various of these irrelevant factors, correlations remained low. Hall believes that a general learning ability of some importance might be discovered were we able to control the differential influence of motivation and previous practice in the learning situation.

On examining both investigations critically the following main points emerge :

(1) The choice of tests in an investigation of this nature is hard to make, and is an important part of the investigation. One must be extremely careful how a given set of correlation coefficients are interpreted, and must investigate the nature of the tasks from which they were obtained carefully.

(2) There is an urgent need for an improved measure of learning ability, as both total score and improvability as measured by the difference between initial and final trials, have very serious drawbacks.

(3) Finding the reliability of a learning task by correlating the sum of the scores found on odd

and even trials, is a procedure which can be severely criticised.

(4) Before accepting any conclusions on the topic of interrelationships between learning abilities, further data is urgently needed. In order to obtain some such data, this investigation was undertaken, the specific aims of which will be set forth in the following chapter.

CHAPTER II

THE AIM OF THE INVESTIGATION, AND DISCUSSION OF THE CONCEPTS INVOLVED.

A. The Problem stated briefly.

The aim of the investigation can briefly be stated as being : to find out to what extent various learning abilities are correlated. In developing a method of investigating the problem, it is first necessary to make a survey of the available literature bearing on the subject. This has been done in the previous chapter. Secondly it is necessary to devise methods of overcoming deficiencies of previous investigations. This includes

(a) Selecting or devising learning tasks specially suited to the problem.

(b) Finding or devising improved methods of measuring learning ability.

(c) Planning the investigation so that it will yield data that can be submitted to statistical analysis.

The rest of the investigation will consist in carrying out the learning experiments on a sufficient number of subjects, to enable one to use the correlation technique (that is well over fifty); and in making a statistical and psychological analysis of the results obtained.

Before selecting or devising learning tasks, or improving previous methods of measuring learning

ability, let us examine some of the concepts used in formulating the problem. How is one to define learning? What kinds of learning are to be considered? What is ability? How can learning ability be measured?

B. The Nature of Learning.

The term learning will first be considered. As used in psychology it seems to cover almost any form of adaptation to a situation which an organism acquires. In spite of the fact that learning rates in different situations appear in many instances to show only slight correlation, many psychologists continue to use the term for all situations, and treat learning as a single function. Maier,(43) however, suggests that there are several separate processes which make up the learning function. To him it seems that the study of learning has reached a point where further analysis is sorely needed. Correlations between different learning problems give one nothing but correlations, he says, and it is time that attention be turned to the functioning mechanisms responsible for the size of the correlations. To the present writer the fact that correlations are low, suggests that the supposition that learning is the resultant of several different processes, may be correct. If learning can be broken up into different mechanisms which constitute the learning process, then a learning situation in which one of the mechanisms is mainly brought into play, may give a high correlation with another learning problem in which the same mechanism is functioning.

McDougall (42) discusses whether learning is all of one type, and comes to the conclusion that there are two distinct types of learning process: intelligent learning involving achievement through relevant insight, foresight and feeling, and unintelligent learning through mere repetition. Concerning this latter, he writes: there is a kind of learning which consists merely in the fixation or facilitation of a series of activities through repetition. He does not maintain that such learning is purely a mechanical process, but holds that it is a conative activity, and if this conative element were lacking, doubts whether mere repetition of a movement sequence would result in facilitation. Spearman (53, page 284) also points out how improvement may be derived from two entirely disparate sources: "first, there is bare retentivity...., and then, there is such improvement as derives from change in mode of operation, and therefore is not explicable by retentivity at all, but solely by eduction".

Koffka (38) seems to make a similar division, when discussing the parts played by achievement and memory in learning. He holds that all learning requires the arousal of configural patterns, and that before such a pattern has been achieved, mere repetition does not aid the learning process. However, after the configuration has once been constructed (achievement), repetition serves to make the behaviour appreciably firmer and easier. Upon repeating the objective conditions, the configuration will arise much more easily and more swiftly than it did the first time (memory). Thus "memory" (Koffka) appears very similar to "facilitation by repetition"(McDougall).

One method of dividing learning into different processes would therefore seem to be into "achievement" and "facilitation" processes.

Facilitation can be regarded as being a retentivity phenomenon. It refers to improvements produced by sheer repetition, as distinct from relation and correlate eduction (Spearman), or configuration construction (Koffka). It falls in line with Spearman's principle of retentivity, about which he writes: "The other way in which the retentivity displays itself is as facilitation: cognitive events by occurring tend to re-occur more easily". From this facilitation in learning is a corollary. (Spearman: 52. Chap.IX.) As McDougall (42) points out, something may be learned so that further repetitions produce no improvement of an observable kind; yet if it is repeated for more trials after this, then after an interval of twelve months, say, it will be found to be better remembered, than if it had not been repeated in this manner. Therefore there must have been some improvement in retention due to the extra repetitions. This is improvement due to pure facilitation.

Achievement is that type of learning which depends upon the eduction of relations and correlates (Spearman). Reasoning enters into it, and it may be ideational, although in some cases the eduction of relations and correlates probably takes place subconsciously. Koffka (38) points out how the problem of learning has often been identified with the problem of memory, and emphasises the importance of achievement, which is generally the process whereby the first

performance comes about, especially in learning of the problem-solving type. To him this involves the power to form the required configuration to solve the problem. In learning of the habituation type, he holds the function of repetition is to prepare the ground for the construction of an appropriate figure, which first occurs by chance. It is only after the configuration has once been constructed, that repetition serves to make the behaviour firmer and easier.

One may not agree with all the theoretical implications in the discussions of Koffka, McDougall, or Spearman, but all three authors regard achievement and facilitation as two very different processes, which occur in the learning situation, and this is the point the writer wishes to stress. Leeper (40) in discussing future work on learning which he would like to see undertaken, also says that it would be desirable to make a division between the different phases of learning. The acquisition phase of learning (as distinct from utilization) needs to be recognised as composed partly of <u>discovery of the solution</u> (compare with achievement), and partly of the <u>fixation</u> or implanting of the discovered solution (compare with facilitation). In all actual instances of learning the two processes, facilitation and achievement, are probably intimately blended in various proportions. In many learning situations, at the beginning achievement is chiefly responsible for producing the improvement; in others the parts played by achievement and facilitation may be equally important, while in some perhaps facilitation is the important factor. In all learning there is the acquisition of a new mode of reaction or

adjustment (achievement) and also the retaining of it. (facilitation).

C. <u>The Choice of Tests according to the analysis of Learning into Facilitation and Achievement Processes</u>.

Were the low correlation coefficients, reported by previous investigators, perhaps due to the inclusion of these two very different processes in varying amounts, in the learning required by the tasks they used? In order to find out whether this were so, it was decided to try and choose eight tests, four of which, as far as possible, would involve facilitation prominently in the learning process required, the other four involving achievement in a more marked degree. (Achievement meaning the opportunity to educe relations and correlates, use reasoning, etc.) If the two groups of tests could be kept very similar, apart from the amount of facilitation and achievement that could be used in the respective sets, this would make the correlations found easier to interpret. For example, if a fairly high correlation were found between the results obtained from the various problems devised as facilitation learning tasks, and also a high correlation between the achievement tasks, but a low correlation between similar tasks in the two groups, then the reason for the sizes of these correlations, would be that different processes are required to make a good facilitation learner, and a good achievement learner.

In an investigation such as this, where the eight tests chosen must be easy to compare with each other, it was found difficult to devise tests in which

the proportion of achievement, or facilitation is great in one group of tests, and not in the other. If four tests, where the process required for learning is chiefly achievement, are chosen, these would give curves of the typical puzzle-box variety, that is, a long initial time, followed by a sudden drop and flattening out of the learning curve. In regard to measurement of improvement such curves would be difficult to compare with ones in which the improvement is steadier and more gradual; that is, the type of curve one would expect to obtain from a task involving mainly facilitation. In making two sets of tests, very similar for the reasons stated above, the investigator found that the amounts of achievement and facilitation involved in both sets also tended to become fairly similar. In the end, two sets were decided upon in which a greater opportunity for using achievement learning was offered by the second group of tests, but the one set does not involve purely facilitation processes and the other only achievement as we would have wished; the difference between the two sets is only one of degree.

Another consideration is that due to the similarity of the tests, some subjects may in learning the second batch of tasks, designed as achievement tasks, actually use mainly the facilitation method, while others may as intended make use of the opportunities afforded for utilising achievement in the learning. Seashore (51) points out that one cannot simply control amount of training, and attribute all else to biological capacities, as the work method employed is a third variable. Therefore it is necessary to word

the instructions very carefully, to try and ensure that subjects will use the achievement methods of learning as much as possible in the second group of tests. Dashiell points out that subjects often change "set" in spite of instructions, therefore one cannot be sure that subjects are using the achievement method. They may also alternate between the two methods. (13).

D. The Concept of Learning Ability.

Having discussed learning we now come on to a consideration of "learning ability". The term ability is one which is coming to be used more and more frequently in psychology, since the old faculty psychology has fallen into disfavour. When we speak of a person having an ability, we infer the existence of it from an observed act. However the ability is something different from the act. The act is something which the individual does, whereas the ability is something which the individual possesses. When one speaks of an individual having a certain ability, for example aiming ability, one does not imply that the person did well in an aiming experiment on one specific occasion, but the term ability would seem to refer to the possibility of repeating the act on a subsequent date. If the act can be repeated in this way, its occurrence must be dependent on something which endures for some time, therefore ability would seem to be dependent on the constitutional nature of the organism. Ability seems to define the reactive nature of the individual regarding his possession of the constitutional factors upon which the occurrence of the act is dependent, both now and in the future.

A person can be said to possess the ability to do whatever he is observed to do. The chief question of psychological interest is the relative amount of ability possessed by different individuals, and any judgment of amount of ability is necessarily based upon some quantitative feature of the corresponding act. The relative amounts of ability possessed by the members of a group of individuals are indirectly measured by means of the distribution of efficiency scores that is known to be a functionof the constitutional differences of these individuals.

Learning ability is distinctive from other abilities, in that it is a developmental concept, as contrasted with other abilities which may be called "Performance Abilities". Learning ability differs from these in respect to the goal which the individuals are striving to attain when the ability expresses itself, and in terms of which their efficiency is measured. The goal of learning is future efficiency, not present performance, therefore learning ability is a developmental concept. Thus one may have a certain ability now in dart throwing (performance ability) but one also has an ability to improve in dart throwing with practice (learning ability). Hence learning ability can be defined as the power to improve on one's performance of a task which is repeated a number of times. Guilford (21) writes that certain studies lead one to the conclusion that practice in tests does not augment the fundamental abilities, but changes their relative importance in performing those tasks. In this case learning

ability would be the ability to change the relative importance in a task of the various fundamental abilities. For the present purpose the power to improve in the performance of a given task seems the best definition of learning ability as this investigation is not particularly concerned with the way in which this improvement occurs.

Learning ability and intelligence quotient (I.Q.) are concepts which have been closely related. I.Q. is often defined as a measure of learning ability, a measure of intelligence, and a measure of mental growth (Carr & Kingsbury, 7.) This definition implies that a general learning ability exists. Guilford points out that recent investigations should be sufficient to do serious damage to the definition which identifies intelligence with learning ability. If learning abilities do not correlate highly with each other, it is not likely that they will correlate with I.Q. Evidence on this point is given in Husband's investigation.(30). The range of correlation coefficients between the various tests and intelligence is from -.16 to +.29, with a median coefficient of .005. Seven of the coefficients are negative, and seven positive. Considering these figures it would seem dangerous to identify learning ability with intelligence.

In the realm of cognitive abilities Spearman (53) has revolutionised our thinking on the subject by his two-factor theory. He maintains that a certain amount of g (or the general factor common to all performance abilities) is necessary for success in every cognitive activity, which is

not so purely habitual or innate as to proceed unconsciously. However in every distinct ability there is also a specific factor or factors, S or S's. Activities which resemble each other very closely indeed, usually have some of their S's in common. There has been a tendency recently to refer to these examples of s-overlap as the "group factors" of ability. In order to find out whether certain abilities are divisible into g and independent S's, or to find out whether s-overlap occurs, factorial analysis is applied to the table of intercorrelations obtained between the various tests. However the time does not seem ripe for the application of factorial analysis to the intercorrelations found between learning tasks, as the measures of learning ability need to be purified before the results of factorial analysis can be satisfactorily interpreted. As will be pointed out in the next section, the scores made on a learning task are contaminated by many variables extraneous to the learning process. Therefore we cannot be sure that the relations between learning tasks are solely due to learning ability, and therefore not much can be gained by applying factorial analysis at this stage.

From this discussion it can be seen that ability, and learning ability in particular, is an important concept, both for systematic and educational psychology, and both these departments of psychology are in need of data concerning the nature of learning ability, the main problem being to find out whether a general learning ability exists, whether learning abilities are specific to the material being learned,

or whether there are group learning abilities.

E. Measuring Learning Ability.

As has already been pointed out any measurement of amount of ability is necessarily based upon some quantitative feature of the corresponding act. But an efficiency score obtained from a given performance is a function of many other factors than ability, the chief of these being objective conditions, volitional attitude and motivation of the individual, fluctuating mental and organic conditions and chance factors (Carr and Kingsbury, 7.). When dealing with the learning situation the case is further complicated by the fact that an individual's scores are influenced by previous practice, positive and negative transfer of training, the number of trials given on the specific tasks (Hall 23), and work methods (Seashore 51; and van Dusen, 16).

As a first step in measuring ability the experimenter must secure a group of scores under such conditions that their distribution cannot be explained in terms of any of these factors, but only in terms of constitutional differences. The difficulties in the way of doing this with regard to learning ability are made clear by Hall (discussed in Chapter I). An attempt has been made in this investigation to overcome at least some of these difficulties, and to study the effects of eliminating some extraneous factors, on the intercorrelations found between various learning abilities.

Due to the fact that learning ability is a developmental concept, one must also decide how to

measure it from the set of scores obtained from a number of trials on some learning material. Learning ability is the power to improve on the time taken, on the number of errors made, or on the amount done in a given time, in a given task. Most people have the power to improve in this way when repeatedly performing a task. To guage a person's learning ability, one must find some means of measuring the increase in efficiency. It is difficult to know what the best measure of improvement is. The following methods have been used in the past: the absolute gain method (found by subtracting final from initial score); the percentage improvement (found by representing final score as a percentage of initial score), common points of mastery technique (taking either the number of trials required to improve from the initial common point, to the final common point chosen; or amount of gain which occurs in a certain number of trials after the subject has passed through the initial common point of mastery chosen); rate of improvement (represented by the gradient of the curve); and total time taken to do a certain number of trials (which involves the assumption that everyone starts with about equal ability on the task). It was found that each of these methods had its drawbacks and could be criticised on various grounds. A discussion of these methods will be found in Chapter VI.

It was felt by the present investigator that one must take as many aspects as possible of the learning curve obtained, into consideration, before trying to estimate a person's learning ability from it. A part of this investigation has therefore

been devoted to trying to find some formula which will serve as a better means of measuring the degree of learning ability which the individuals tested exhibit on various learning tasks.

F. Summary.

The aim of the investigation can now be restated: to find out to what extent various learning abilities are correlated. Having discussed the meaning of such concepts as "learning" and "ability", the writer can be more explicit as to how she intends carrying out the investigation.

As there seems to be a growing feeling among theorists, that the learning process is made up of different mechanisms, it was felt that the low correlations reported by previous investigators may have been due to the inclusion of the "facilitation" and "achievement" processes in varying amounts, in the different learning tasks they used. In order to find out whether this is the reason for the low correlations it was decided to choose eight tests, four of which, as far as possible involve facilitation prominently in the learning process required, and the other four involving achievement in a more marked degree. The two groups of tests are intended to be as similar as possible,apart from the fact that a greater amount of achievement can be used in one set, in order to make the correlations found easier to interpret.

From the discussion of "ability" and "learning ability", the importance of aiming at the construction of learning tasks in which the scores

obtained would be as free as possible from the influence of objective conditions, volitional attitude and motivation, fluctuating mental and organic conditions, previous practice, positive and negative transfer, and work methods, was realised. It is also felt to be necessary to find an improved method of measuring learning ability from the set of scores obtained by giving an individual a number of trials on a learning task.

In our discussion of ability, the possibility of applying factorial analysis to the intercorrelations between learning tasks was considered, but it was decided that before this could be done with profit, it must be ensured that the scores obtained on the learning tasks are not contaminated by variables extraneous to the learning process.

The final solution of these problems will require the application of factorial analysis, but until that is done, it must be remembered that relatively low intercorrelations do not forthwith prove the absence of group or general factors. Once the tasks have been selected, and a formula for measuring learning ability has been devised, the investigation will consist of giving the learning tasks to about seventy subjects individually, applying the "learning ability formula" to the results, and then finding the intercorrelations between the tasks. The final part of the investigation will entail a psychological analysis of the results, for as Bartlett (2, pg.8) so strongly emphasises, if statistical applications in the field of psychology are to have any value whatsoever, they must be both preceded by & also supplemented by observation and interpretation.

CHAPTER III.

THE CHOICE OF LEARNING TASKS.

A. Division into "Facilitation" and "Achievement" Tasks.

The previous chapter dealt with the theoretical considerations which led to the decision to choose the learning tasks in two groups, one group to be comprised of four tasks involving, as much as possible, facilitation in the learning process by which the task can be mastered; the other to contain four tests, as similar to the tasks in the first group as possible, except for the introduction of certain elements which would allow of more achievement learning. It may be argued that if evidence for general learning ability is being sought, and if high correlations are found by this method, this may be due to the similarity of the tasks chosen. However one task is similar to only one other, and if these are the only two tasks which correlate highly with each other, then obviously a case cannot be made from this for a general learning ability. High correlations between disparate tests are being sought, and therefore the four individual tasks within each group were made as different from each other as possible.

The first division within each group was into "motor", and, what for lack of a better term the writer has called "mental" tasks. Two motor and two mental tasks were chosen in each group, the type of operations involved in the motor tasks and mental tasks being as different as possible. If any correlation is found between the different tasks

in the facilitation group, say, it must be assured that this correlation is not merely due to the similarity of the learning tasks chosen for that section.

Considering these points a scheme was drawn up into which the tasks chosen had to fit before being deemed suitable.

A. Facilitation Tasks.	B. Achievement Tasks. (Achievement Tasks being as similar to the facilitation tasks as possible, except for the fact that they offer more opportunity for educing relations and correlates).
1) A motor task.	1) Motor task similar to A 1), giving an opportunity for relation eduction.
2) Another motor task, involving as different as possible a type of operation from A 1).	2) Motor task similar to A 2).
3) A mental task.	3) Mental task similar to A 3).
4) Another mental task, involving a different type of operation from A 3).	4) Mental task similar to A 4).

A scheme for the choice of Learning Tasks.

Four facilitation tasks were then sought, in which no "system" was present as an aid to learning, so that the learning would be largely dependent on retention, obtained by repeating the task, until such retention was acquired.

In the achievement tasks an attempt was made to introduce a definite but fairly complex "system" into the material to be learned. The learning would

then depend on relation and correlate eduction (plus retention). The learning however should be gradual, therefore the system was not to be of such a nature that once it had been detected, the task could almost immediately be perfectly performed. If this had occurred, curves of the type indicated in Figure 1 would have been obtained - a curve in which the initial time was long, but once this had been completed the subsequent times were near the physiological limit.

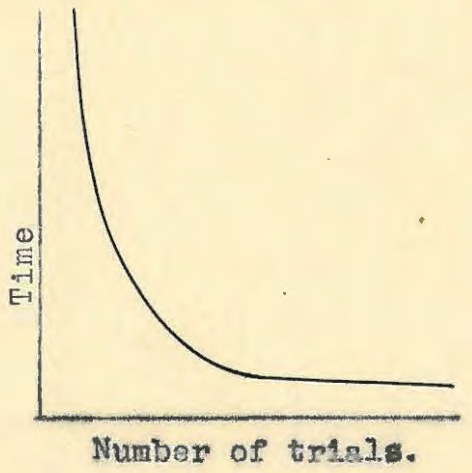

FIG. 1. SHOWING TYPICAL LEARNING CURVE FOR PROBLEM-BOX MATERIAL.

As has already been pointed out in the previous chapter, this type of curve would be very hard to compare with the type of curve expected from the facilitation tasks, where the times would be reduced gradually - a bit more rapidly at the beginning than the end, but still gradually. As Hall (23) points out that the first trials of almost any learning tasks are notoriously unreliable, it was decided to ignore the scores obtained on the initial trials in this investigation. If the initial time were ignored in a curve, such as shown in Fig. 1, the curve would become meaningless, and could not be compared with a more gradual type of curve. It was therefore felt to be important that the achievement tasks be

ones which could only be learnt gradually. This meant that perhaps more than one "system" had to be introduced into a task, so that the subjects' time would not drop too suddenly, once the system had been educed.

The experimenter tried to guard against the possibility of the achievement tasks being done by the facilitation method, by pointing out that there was a system, and that if this could be educed the improvement that would show up would be greater than otherwise. As the facilitation and achievement tasks were being kept as similar as possible, it was found to be almost impossible to devise tasks that must perforce be performed by the achievement method.

B. Tests all measurable in Time Units.

The tests were all chosen so as to be measurable in terms of the same unit, as it was felt that the low correlation coefficients found by previous investigators, may have been partly due to differences in the unit of measurement used for the different tests; for example, decrease in the number of errors may have been taken as the indication of learning in one task, and in another time may have been used as the unit of measurement. The correlation between these two tasks may be lower than that found by using the same unit of measurement for both tests.

It was considered that time would be a suitable unit, as tasks could be devised, in which, by merely lengthening or shortening the amount to be done in each trial, the time taken for a trial on one task could be directly compared with the time taken on a similar trial, for another task. This could not be done so easily using errors, or amount done

per trial as the unit of measurement.

Time affords a finer measure of the learning going on, than does number of errors, or even amount done per trial. One can see how the gradations are finer, using time rather than errors, if one considers a hypothetical subject's performance on a maze. She may go into eight blind alleys on three successive trials, but the time taken for those three trials may be 34 seconds, 29 seconds and 25 seconds, which shows that although actual errors were not being eliminated, improvement must have been taking place in the subject's method of running the maze, which should surely be accounted for, in the measure used, and time is the measure which would account for this, while number of errors would not.

There are certain objections to using time as a unit of measurement. A subject may take such a long time over a certain trial that for purposes of comparison it is hard to know how to deal with her; whereas if number of errors is the unit to be taken, a subject cannot make more errors than the nature of the test material affords. This applies particularly to mazes. Also, if animals are being used as subjects, time is a very unreliable measure to take, as the animals may curl up and go to sleep during a trial, and time taken would be no indication of its knowledge of the maze.

The amount done in a certain time is only applicable to certain kinds of material and could not very well be used in such a task as a maze. Even in test material, such as a form-board, time taken for a complete trial is a better unit to use than amount done in a certain time, as certain pieces may be more difficult to fit in than others, and therefore a

record of the number of pieces fitted in, may not always be a true record of the improvement which has taken place.

Using time as the unit of measurement meant that all the tests had to be given as individual tests. From the point of view of amount of time taken for testing, this is a drawback. In this investigation two hundred and ten hours of testing had to be done. It may also be argued that conditions cannot be kept as constant for a large number of people as they could if the tests were given to that number as a group. This argument can be outweighed by the fact that people respond very differently to the group situation, and, while everyone is being treated in objectively the same way, this way may stimulate some people to react to the best of their ability, while others may become over-motivated and not show their true amount of learning ability, due to emotional factors. In the individual testing situation the experimenter can often do much to calm down an over-excited subject, and without the presence of other people doing the same task, excessive conation does not so often occur, and prove an interfering factor.

C. Choice of Tests.

Many tests were considered and tried out on a few subjects, before eight tests, which seemed suitable, were chosen. Much difficulty was experienced in getting tests that were not too difficult, and yet not too easy; tests in which sufficient improvement occurred for most of the subjects, over ten trials, and tests which enabled a subject to do ten trials on two tests, within a fortyfive minute period. In choosing the facilitation tests, a consideration

which influenced the choice a great deal, was that the tests chosen had to be convertible into tasks involving more achievement by the introduction of a system that would make learning possible by the eduction of relations and correlates, and not so dependent on repetition.

a) Motor task involving facilitation.

Following the plan on page 40, the investigator first tried to find a suitable motor task, involving mainly facilitation. It was thought that tracing a simple figure, perhaps blindfolded would be suitable. The subject could be given a board in which a groove, such as shown in Fig. 2, had been out, and be required to trace this with a stylus in the shortest possible time.

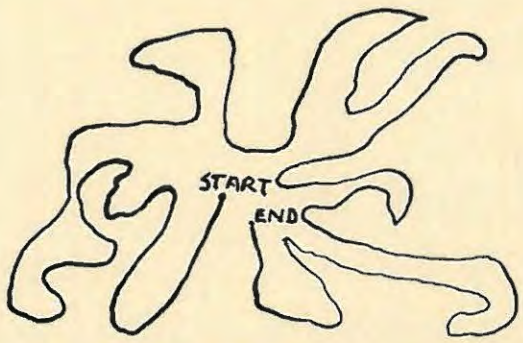

FIG. 2. SIMPLE GROOVE FIGURE.

There are no blind alleys into which the subject can go, therefore improvement will be due to increased facilitation in tracing the figure acquired by mere repetition. This appeared to be an admirable test in which facilitation would have been mainly involved in the learning process required, achievement probably only playing a part in the first trial. It was felt however that the task was of such a simple nature that improvement would not show up over ten or fifteen trials.

From this idea of a simple figure groove, evolved the idea of having a maze. The maze had to be

one which did not allow of much spatial relation eduction therefore a linear maze was decided upon. The maze took the form of a groove cut in wood, the subject being required to trace the maze blindfolded with a stylus. In this task the subject had to learn a sequence of movements, in which there was little spatial relationship to be educed. Furthermore there was no system such as two moves to the right, followed by two moves to the left which would help the subject in learning.

The unit of measurement would be time, and improvement showed up until the end of fifteen trials in the subjects upon whom this test was tried out.

b) Second Motor Task, involving facilitation.

For the second motor task, in order to get a test involving a different type of operation from maze tracing, a task involving more manipulation was chosen.

In this connection marble sorting, binding Meccano bars together with screws and nuts, the Witmer cylinder test, and the Cornell form-board were all considered and tried out on a few subjects. None of these seemed suitable as the tests either offered too great an opportunity for relation eduction or else improvement ceased to show up after the third or fourth trial.

One of the tests used by Husband was spool-packing, in which the subject picked up one spool with each hand and placed them on a tray. Twelve spools filled one tray and five trays constituted one trial. Husband only gave his subjects five trials and it was felt that if ten trials were given improvement might cease as in marble-sorting. Spool-packing suggested block-packing, and the block-packing section of the

Detroit Manual Ability test was tried as a learning experiment, the subject packing the blocks blindfolded and using her left hand. Although a fairly steady improvement showed up, it was very slight. The idea of having different sized blocks then occurred to the writer, the blocks always to be placed in the same position on a shallow tray. Achievement would probably play quite a large part in this test during the first couple of trials, but once the positions had been memorised the improvement would be due to the facilitation taking place in repeating a sequence of movements in a certain order.

c) A Mental Task-Facilitation.

Learning some meaningless material by mere repetition seemed the obvious test to give for a facilitation learning task involving mental activity. Nonsense syllables appeared to be useful material of this meaningless type. Learning a list of nonsense syllables by repeating the list a number of times would not give a curve, with time as the measure of improvement. In order to convert the material into a task, such that time could be used as the measure of learning, a mental maze was devised. The syllables were arranged in sets of four, one syllable in each set being the correct one, which the subject would learn by repetition. The form of presentation and the form of reaction required of the subject proved to be problems. Visual presentation on typed sheets was finally decided upon. As regards the subject's reactions, a board arrangement was considered in which holes were bored, four holes representing the four syllables in each set, the subject being required to stick a stylus into the

hole corresponding to the position of the correct syllable. If this were done an electric circuit would be completed and a bell would ring. As a task involving manipulative ability was to be avoided it was decided that this scheme merely gave the test a pseudo-scientific air and that the scheme of letting the subject read the syllables aloud, the experimenter saying "Yes" after the correct syllable in each line had been read, be adopted. The criticism of this procedure which may be put forward is that the subject will increase her reading-rate, so that the times recorded will not be a true record of the learning taking place. However in learning a maze the subject improves in her manipulation of the stylus round paths, as well as in her knowledge of the direction of the correct path, therefore increase in reading rate in a nonsense syllable maze can be compared with increase in rate of moving a stylus in a groove maze.

In order to obviate the possibility of the subject merely learning the place location of the syllables and not the actual syllable correct in each line, it was decided that the syllables should be differently arranged in each line from trial to trial. Further details as to how the test was drawn up will be found in the next chapter.

d) <u>Second Mental Task - Facilitation.</u>

For this purpose a substitution test was chosen, as in such a test a subject learns which numbers represent which letters, and does so mainly through repeatedly writing them down. Of course a substitution test involves more than this type of facilitation, as at the beginning of the test the

subject is mainly learning the positions of the letters and numbers, so that he can refer to them quickly without looking right along the line of the key, each time he wishes to find a given letter. Other types of mental facilitation tasks seemed too similar to the mental maze test and therefore a substitution test was decided upon.

e) A Motor Task, involving Achievement.

To correspond with the linear maze, a maze having more spatial relationships was chosen. The passages were of varying lengths and the general direction of the correct path changed considerably, so that spatial and directional relationships may be educed, and the eduction of these relationships would mainly constitute the learning required.

f) A Second Motor Task, involving Achievement.

For the second motor task on the achievement side, it was decided not to have a task too similar to the block-packing, as all the other tasks on both sides corresponded very closely. The Cornell Form-board and Mirror-drawing were both considered, but it seemed very hard to decide whether the amount of achievement possible in these two tests was any greater than that utilised in learning to pack blocks, and therefore these tests did not seem suitable.

A card-sorting experiment, as suggested by Woodworth (61, pg. 160), in which the cards have to be sorted into pigeon-holes, was considered as a suitable motor task involving achievement. The arrangement of the numbers of the pigeon-holes should have a system

and the cards should always be stacked in a definite order. This task involves a somewhat similar manipulative ability to block-packing; the improvement in the card-sorting need not be mainly dependent on the improvement in manipulative ability and memory of positions, as in block-packing, but could be due to educing the relationships afforded by the arrangement of the numbers and the cards.

g) <u>A Mental Task, involving Achievement.</u>

Corresponding to the nonsense syllable mental maze, it was quite easy to devise a similar maze, in which a system whereby the correct syllables are to be chosen was introduced - for example, every correct syllable could start with the same consonant. In order that the improvement would not be too sudden once the system had been educed, it was decided that a different system for each set of four lines should be introduced.

h) <u>A Second Mental Task, involving Achievement.</u>

Corresponding to the first substitution test another substitution test was devised with a system that could be educed. It was difficult to find one system long enough to cover more than nine figures, and it was felt that if only nine figures were used, these might be learnt by the facilitation method before the system had been detected. It was again decided to introduce different systems, four figures having one system, the next four another, and so on. The systems which were finally devised for 18 figures, which were not exactly "systems" in the truest sense of

the word, but the shape, sound or "number of syzygy" associations that could be formed between the letters and figures, were made as obvious as possible.

In the next chapter a detailed description of the best material used will be found.

CHAPTER IV.

DESCRIPTION OF TESTS.

A. The Actual Tests.

The following is the set of eight tests finally decided upon:

Facilitation Tasks.

Test I. Linear Maze (Maze A).
Test II. Block-packing Test.
Test III. Substitution Test (A).
Test IV. Nonsense Syllable Maze (A).

Achievement Tasks.

Test V. Maze (B).
Test VI. Card-sorting Test.
Test VII. Substitution Test (B).
Test VIII. Nonsense Syllable Maze (B).

A detailed description of each test follows.

TEST I. The Linear Maze.

This consisted of a groove in shape and size exactly the same as the diagram shown in Fig. 3. cut out of a rectangular piece of three-ply wood, 15" x 8" x .2". This piece of wood was then glued to a sheet of glass to ensure that the floor of the maze would be smooth, and that grooves would not develop in the correct path as often happens in mazes having wooden floors.

The "stylus" with which the subjects were required to trace the maze consisted of a wooden pen-holder, to be held upside-down. The final half inch

of the penholder had been whittled and sand-papered to make it cylindrical, being approximately .1" in diameter.

When the subject entered the testing room, after some preparatory remarks (see Chapter V), the following instructions were read to the subject:

"The maze consists of a grooved path cut into a piece of wood. It consists of a series of forward moves, interspersed with turns to the right and to the left. You are required to find your way through the maze, blindfolded, in the shortest possible time, by pushing the stylus along the groove. By keeping to one side of the groove you might eventually get through, but you will reduce your time most if you learn to avoid the blind alleys and follow the correct path.

Do not be disturbed if the stylus makes a squeaking noise, this is only because it is moving on a glass surface.

Hold the stylus more or less vertical with your fingers round the green part. Do not let your hand touch the actual maze.

You will be given a number of trials, in order to see how much you can reduce your time with practice. I will start you off by saying "Ready - GO", but do not start until you have heard GO. When you come to the end I shall say "Right".

The subject was then blindfolded, and required to trace the maze ten times. The time taken for doing each trial was recorded, as also the number of entrances into blind alleys per trial. The squeaking of the stylus was eliminated by constant sandpapering of the tip, so that it did not become too smooth.

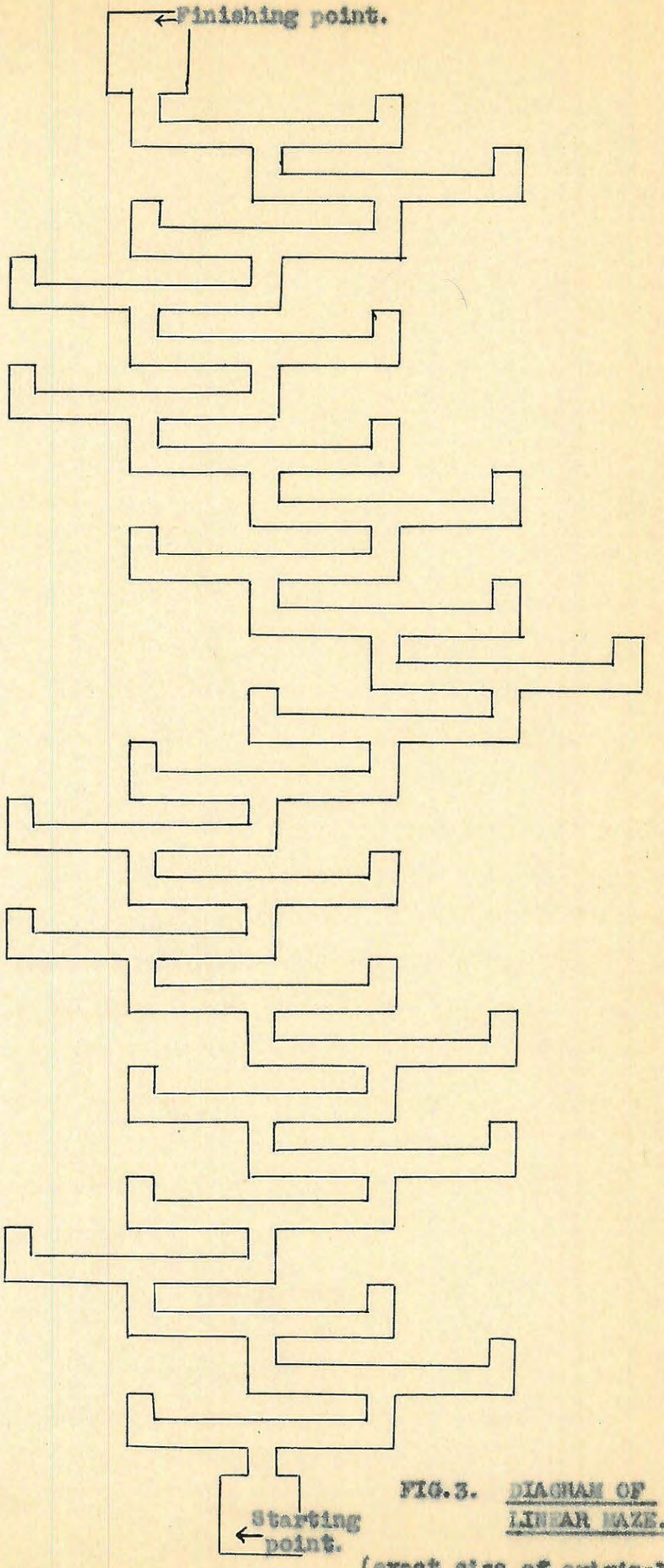

FIG.3. DIAGRAM OF LINEAR MAZE.

(exact size of original).

TEST II. Block Packing Test.

The apparatus for this test consisted of a wooden board, 9" square, around the sides of which were nailed strips of wood, the upper and lower strips being 1" x 7" x .2", and the side pieces being 1" x 9" x .2", so that a form-board frame, 7" square and .2" in depth, was formed. Into this frame fitted rectangular blocks in shape, size and arrangement corresponding to the diagram shown in Figure 4.

The subject, before doing the test, is asked if she is left or right handed. The following instructions are then read:

"You will be given a board, in which there is a square depression of this size (showing plan (Figure 5) of a similar form-board), into which sixteen rectangular blocks are to be packed. They are not to be packed in exactly the same way as in the diagram I have shown you - that is merely to give you an idea of what to do. I will give you the blocks one at a time, always in the same order. When you have put the block in its correct position I will say "Yes". The blocks must be packed approximately from the right to the left-hand side of the box. Each block must always be put in the same position as it was on the first trial. If you are right-handed pack with your left hand, keeping your right hand over the blocks already packed to prevent them from becoming displaced.

I will place each block just here (indicating spot on the table) with my hand just behind it, so that your hand will be stopped if it moves too

much towards the left.

You will be given a number of trials to see how much you can reduce your time with practice. I will start you in the usual way by saying "Ready - GO".

If the subject is left-handed, the blocks are placed on the right-hand side, and she is required to pack with her right hand.

The subject is then blindfolded, and allowed to handle the board for a few seconds. Ten trials are given in the manner outlined in the instructions, the time taken to place all the blocks in their correct positions on the board being recorded for each trial.

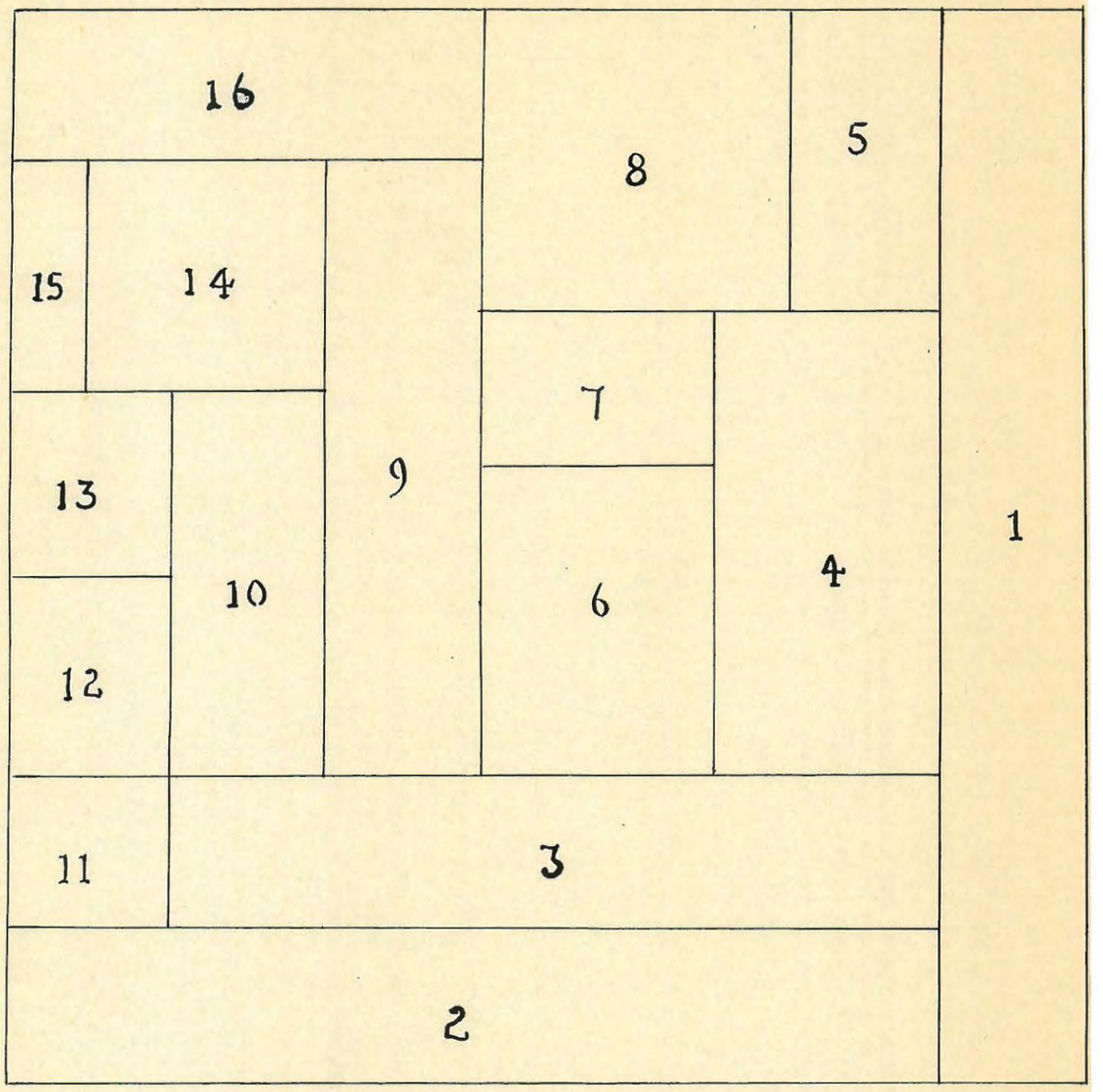

FIG. 4. SHOWING SHAPE, SIZE AND ORDER OF ARRANGEMENTS OF THE BLOCKS IN THE BLOCK-PACKING TEST.

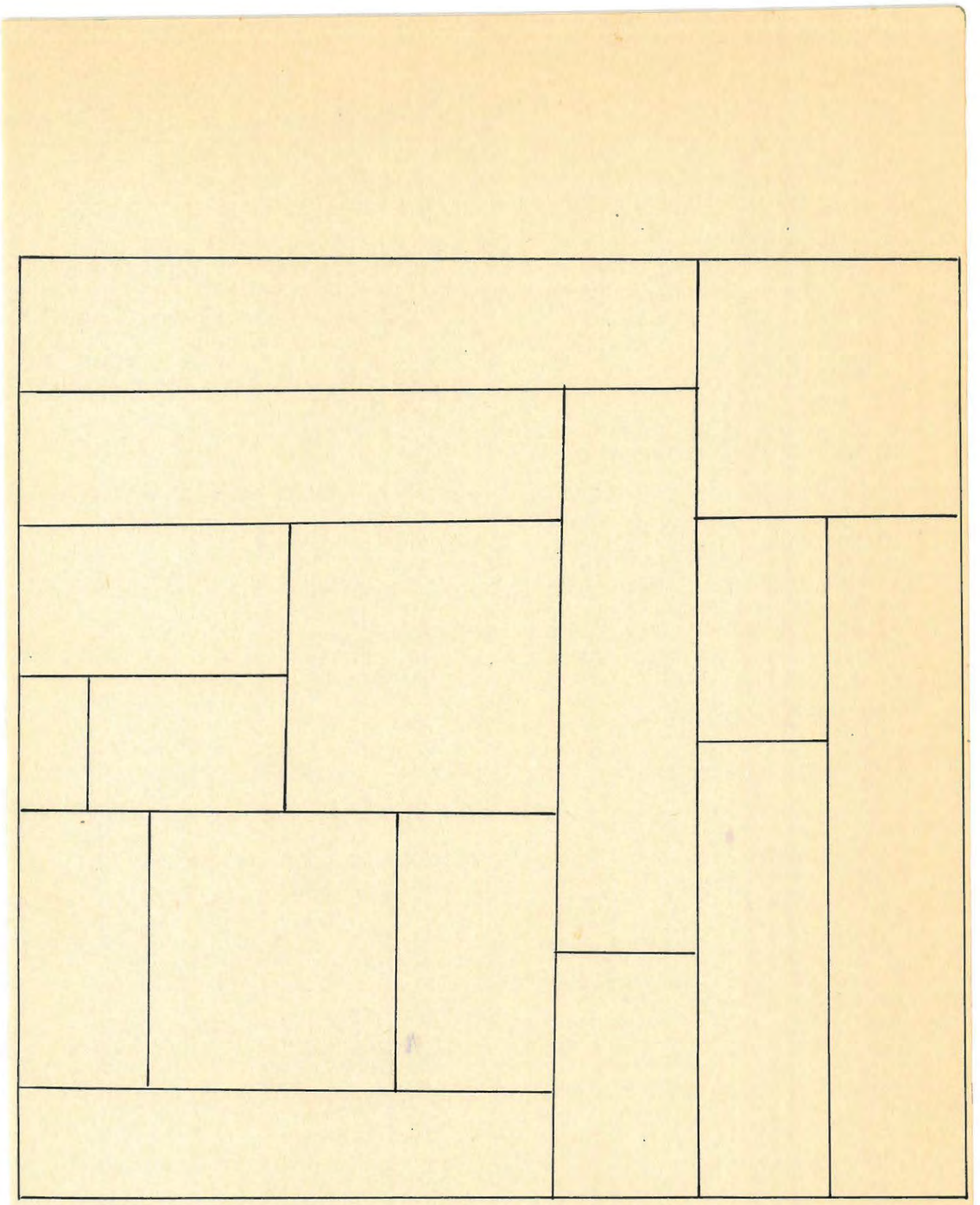

FIG. 5. DIAGRAM SHOWN TO THE SUBJECTS AT POINT INDICATED IN THE INSTRUCTIONS.

SUBSTITUTION TEST A.

A W J B R H T V Z G L K S O H I Q P B H G O N
W T R P G I S L J V A Z H L P R H T A Q H Z W
K V Q O V B J H T K I O A G H L Q S R K B W I

J P Z Q H W S R K A T G H B L O P H Q V Z I J
H O I K P R L J B W Z G L P A R J I G H P T Z
V A L Q G I A B L H R S H K V Z L J O X Q S R

T W J K H O S R I G B V Q T Z W H B L J O K Q
S R T W B H A J S P T H R I K A V Z H G J H P
T Z Q A W O G W H I Q B S V T H O K A R L W H

S G K P V H I B O T Z J A O P S V B I K Q L B
Z V H A J Q O R S G T P W L I G Q R A W O H K
L P V T S J H A H Z Q I H B R W Z T J L S P H

O P V Z W R Q O G J L H T P I V R H P G K Q I
S W T A V K O W B V H B G R O H Z L H Q J I P
V A G K P W S J R I Q L H B A W T Z V G S Q R

R H B O K I R Z J A K P T V G H Q B O H P L G
S W T K A R I O H Z L I H Q J T Z S P B L J H
O R P S V A W P O T S H R W G H L H K T P A B

J I R Z H W B H A O P L G H V S R W O H H S T
B H K I H Q H V W J V K Q A L B V Z I K G J P
H A B J T Z R G P G Q B A Z T I J L K H Q B W

Z I L S R J O H R T V Z W Q O K H G S H B K A
Q L K I V T J I S R Q W B A H O L Z G H A V W
O R P J A I L P Q J H R T V Z S Q O P L K B H

G I J T Z B W G V I T J S K Q L O P A R S P G
V Q T H V A Z P H P I R L I J K R G S W Z H G
I R Q W T L K H I T J H Z G P L H W Q B A R T

TEST III. Substitution Test A.

The appended sheet of letters was given to each subject. This sheet of letters was drawn up so that each letter should occur approximately the same number of times in each trial (i.e. set of three lines). This was done by plotting the letters into squares (23 squares across and 3 down). In this arrangement each letter should occur four times in each trial, bar three letters. Each letter, taken at random was plotted haphazardly in the squares, each letter being plotted four times, except for three.

The following key was typed on white paper and pasted onto a piece of cardboard 8.2" x 3.5".

G X S P O V B L W Z N E Q T A R I J.
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18.

This cardboard was moved down after each trial so that the key was always the same distance away from the letters during each trial.

The following instructions were read to each subject :

"You have before you a list of letters of the alphabet. Below each letter write the corresponding number from the key - e.g. A, W, J - (15 9 18) As you proceed you will learn the numbers corresponding to each letter, and so will be able gradually to make the substitutions without referring to the key. This learning will show itself in greater speed. I will take the time you require to do three lines of letters. I will give you the usual signal "Ready - GO", and you will stop at the end of each three lines. After a short rest I will

give you the signal to start the next three lines and so on.

You must work steadily across the page from left to right."

Ten trials were given, the time taken for each trial being recorded.

TEST IV. Nonsense Syllable Maze A.

This test consists of eleven pages of nonsense syllables, two of which are included here as specimens. These are dealt with as described in the instructions, which are :

"I will give you a sheet of paper on which are 20 lines, each consisting of 4 nonsense syllables. You must read aloud from left to right the syllables in each set of four. When you have reached and said the correct one, I shall say "Yes", after which you disregard the remainder of that line, proceeding immediately to the next line. On the second and later trials, each line will contain the same syllables as before, but arranged in a different order. The correct syllable in a given line is always the same from trial to trial, but it will occur in different positions in the line. You must try to learn to know the correct syllable for each line. If you remember which syllable in the line is the correct one, say it straight away, without first reading the preceding syllables in the same line. If you make a mistake in thus picking out a syllable, try

the others in turn as before. Work through each page of syllables in this way - obviously the more of the correct syllables you remember, the less time it will take you to work through the list, for you will read out fewer incorrect syllables. In the last few trials you will probably know most of the correct syllables, and be able to pick them out directly.

As usual wait for the signal "Ready - GO"."

The nonsense syllables were obtained from Glaze's lists of nonsense syllables, to be found in the Journal of Genetic Psychology, Vol. 35, page 265. The majority of the syllables were chosen from the 6.67% + .00% "Association Value" columns.

The following is the list of correct syllables :

ZOV; WUK; GIK; KYG; SEB; VUM; NXJ; JUG; QEF; TUV; GAQ; CIJ; DAQ; YOV; ZYT; MIV; QAZ; FYG; JYK; WUT.

The correct syllables were arranged so that they fell in a different position in the line as often as possible. The syllables were arranged in such a way that if the subject merely read from left to right each trial, she would always make 36 errors. In arranging the syllables, the scheme shown in Figure 6 was adhered to, the numbers indicating the position of the correct syllable in the line.

No. of line	No. of Trial.								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	2	3	4	2	4	3	3	2	4
2	3	4	2	4	3	2	2	3	4
3	1	4	3	2	2	3	4	4	3
4	2	3	4	3	2	4	2	4	1
5	3	2	4	4	4	2	3	3	2
6	4	1	3	2	3	4	4	2	2
7	2	4	1	3	4	2	2	4	3
8	4	2	4	1	3	4	3	2	4
9	3	1	2	4	2	3	4	3	2
10	4	3	2	3	1	4	3	2	3
11	4	2	3	4	2	3	1	4	2
12	3	4	2	3	4	1	2	3	4
13	2	3	4	2	3	2	4	1	3
14	1	4	2	3	4	2	3	3	4
15	4	2	3	4	2	3	4	2	1
16	4	3	2	1	3	4	2	4	2
17	2	4	3	2	1	4	3	2	3
18	3	2	4	3	4	1	2	3	4
19	2	3	1	4	3	2	1	4	2
20	3	2	3	2	2	3	3	1	3

FIGURE 6. SCHEME OF POSITIONS OF THE CORRECT SYLLABLES IN EACH LINE, ON DIFFERENT TRIALS.

For trials 10 and 11 the arrangements for 1 and 2 were repeated.

The time taken for each sheet of syllables was recorded.

CEF	ZOV	QIF	XIM
TEV	DAX	WUK	PUB
GIK	ZOM	JEX	VIK
YEK	KYG	HUK	ZYB
PYV	VAF	SEB	QOB
NEG	LEJ	GOK	VUM
KEB	HYJ	MUB	GUG
ZOS	VIB	KIG	JUG
JID	ZUV	QEF	CEJ
YOZ	PYB	NIF	TUV
MEQ	DEF	PAF	GAQ
QOS	ZIL	CLJ	VYB
ZYS	DAQ	QOJ	XYJ
YOV	JYS	GEG	ZUB
KOJ	NUB	EYL	ZYT
RIX	GOQ	WUJ	MIV
ZIQ	QAZ	BOF	JYG
XIV	NAX	FYG	GAX
YEM	JYX	GEF	ZUT
QUB	JIG	WUT	XIL

SHEET I OF NONSENSE SYLLABLE TEST A.

YIM	QIF	ZOV	CEF
DAX	FUB	TEV	WUK
VIK	JEX	ZOM	GIK
HIK	ZXB	KYG	YEK
PYV	SEB	VAF	QOB
VUM	MEG	LEB	GOK
MIB	KEB	GUC	NYJ
KIG	JUC	VIB	ZOS
QEF	JID	CNJ	ZUV
PYB	HIF	TUV	YOZ
PAF	GAQ	BEP	NEQ
VYB	QOS	ZIL	CIJ
QOJ	KYJ	DAQ	ZYS
GEC	JYS	ZUB	YOV
BYL	ZYT	NUB	KOJ
RIX	WUJ	HIV	GOQ
BOF	JYC	ZIQ	QAZ
GAX	FYC	MAX	KIV
YEM	GEF	JYK	ZUT
JIC	WUT	XIL	QUB

SHEET II OF NONSENSE SYLLABLE TEST A.

TEST V. Maze B.

This maze also consists of a groove, in shape and size the same as the diagram shown in Figure 7, cut out of a piece of three-ply wood 8.8" square. This piece of wood was set on glass in the same way as the linear maze, and the same stylus was used for tracing the maze.

The following instructions were given :

"This maze, like the other one you did, consists of a grooved path cut into a piece of wood. It is unlike the other maze in that the general direction of the path is not always away from you, but varies.

You are again required to find your way through the maze, blindfolded, in the shortest possible time by pushing the stylus along the grooves. You will be given ten trials to see how much you improve.

Do not be disturbed if the stylus makes a squeaking noise, and hold the stylus in a more or less vertical position, not allowing your hand to touch the actual maze.

I will start you off by the usual "Ready - GO" signal, and when you come to the end I shall say "Right", after which I shall return your hand to the starting point."

Again the subject is required to do ten trials blindfolded, the time for each trial being recorded.

As well as recording the times, the answers to the following set of questions were recorded for the Maze Test B, the card-sorting test, the substitution

test B and the Nonsense Syllable Test B.

Questions asked and notes made on the Achievement Tests.

1. Do you like doing this test?

Very much indeed	(5 points);
Fairly well	(4 points);
Moderately	(3 points);
Not very much	(2 points);
Not at all	(1 point);
Actively dislike doing it	(0 points).

(a) After trial 1.

(b) After trial 10.

2. Maze B. Do you mind being blindfolded?

Substitution Test B. Do you dislike tests involving figures?

Nonsense Syllable Test B. Do you dislike reading aloud?

Very much	(4 points);
Not very much	(3 points);
Not at all	(2 points);
Find it quite pleasant	(1 point).

(a) After trial 1.

(b) After trial 10.

Cards. Have you had much practice in shuffling and dealing cards?

(The following questions were asked after Trial 10.)

3. (a) Note whether subject gets very upset when she makes a mistake.

(b) Maze B. Does it upset you very much when you make a mistake you weren't expecting to make?

Cards. Does making one mistake upset you and make you make others?

Substitution Test B. Does it upset you very much when you simply cannot think which figure goes with which letter and you do not seem to be able to find them in the key?

Nonsense Syllable Test B. Does getting one syllable wrong upset you and make you forget others that you know?

4. (a) Were you keen on doing well on this test, or didn't you care whether you did well or not at the beginning?
   (b) Did yourattitude change during the test?

5. (a) Note whether subject is especially cautious.
   (b) Were you being extra careful during the first couple of trials?
   (c) Were you a bit nervous at the beginning of this test?

6. Maze and Card Sorting.
   (a) Has this test made you feel tired?
   (b) Were you tired before you started doing the test?

   Substitution Test B and Nonsense Syllable Test B.
   Do you feel that doing the other test first impaired your ability on this test?

7. Maze B.
   (a) Did you find this maze easier to learn than the one you learnt previously?
   (b) What do you think made it different from the last?
   (c) Do you think this maze was easier to learn because there was a fairly definite direction that could be learnt (more spatial relationships)?

   Cards.
   (a) Did you find this test easier to do than the block-packing test?

(b) Did you notice that the cards were always arranged in the same order each trial?

(c) Did any arrangements of the way the numbers were placed help you?

Substitution Test B.

(a) Did you find this substitution test easier to do than the one you did previously?

(b) Do you think this was due to having done a substitution test previously?

(c) Do you think it was due to the fact that it was easier to form associations between these letters and numbers than it was between those presented in the previous key?

(d) Did the fact that I told you there was some sort of system in the way the key was compiled, hinder your learning?

(e) Which systems did you detect?

Nonsense Syllable Test B.

(a) Did you find these nonsense syllables easier to learn than those you learnt previously?

(b) Do you think this was due to having done a nonsense syllable test before?

(c) Was it because the syllables fell into groups, chosen because of certain principles such as initial letters being the same?

(d) Did the fact that I told you that the correct syllables had been chosen according to some principle, hinder your learning at all?

(e) Which systems did you detect?

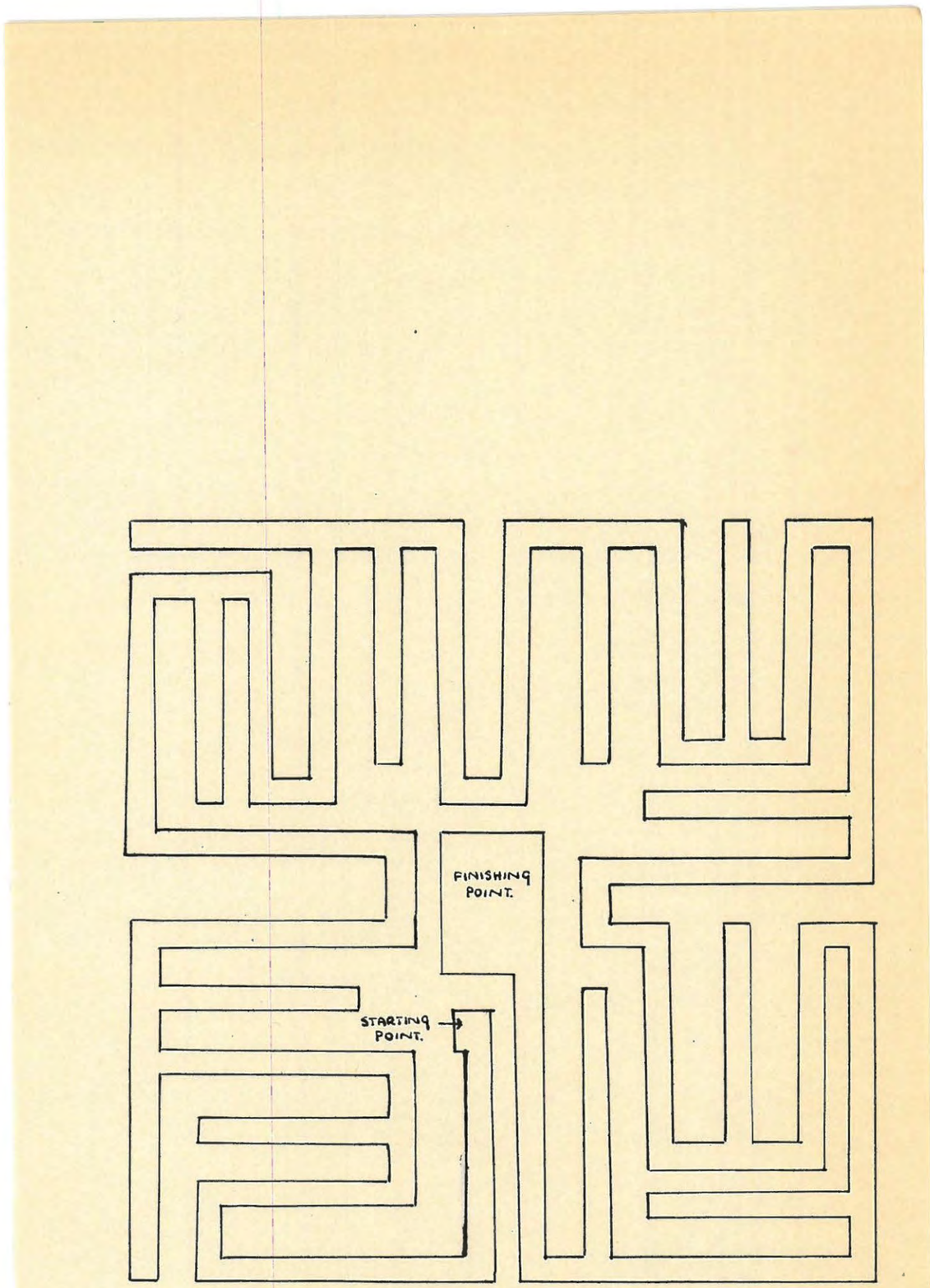

FIGURE 7. DIAGRAM OF MAZE B.

(Exact size of the original).

TEST VI. Card-sorting Test.

This test is adapted from a test devised by Kline and quoted by Woodworth (61 page 160). The apparatus for it consists of a card-sorting box containing 12 pigeon-holes, three rows of four holes each; each pigeon-hole being 4" x 4" x 4".

18	16	14	8
26	65	30	53
10	11	12	13

FIGURE 8. DIAGRAM SHOWING ARRANGEMENT OF NUMBERS FOR THE CARD-SORTING BOX.

The boxes were labelled with black numbers about ¾" high, pasted to the inner side of each pigeon-hole facing the subject, the arrangement of the numbers being that indicated in Figure 8.

The subjects were given a pack of cards from which the aces had been removed and in which there were only three cards of each sort instead of four. The following instructions were then read to the subject.

"This pack of cards is to be sorted into the pigeon-holes of the card-sorting box. Into the top row of pigeon-holes you are required to put the cards which are half the number of that indicated on the box. In the second row you are required to put cards which are the same as the initial digit indicated on the box. Into the third row sort cards 10, Jack, Queen, King, into the pigeon-holes marked 10, 11, 12, 13 respectively.

There are other simplifications in the test,

SUBSTITUTION TEST B.

I T M X H V L Z W S N M A H M Z V R M X A S E
L X K I E B X S T A Z W R V O W R I D O R T Z
K N B M H A S I O K E B T R K B M L W S D X O

M X V L H N Z A R V D I L S D E M X T B Z K W
V S I X T K L O H Z T A D B M K D H W O I R E
D B E R X W S T A O H Z K L E B X V R K H A W

Z H D Z E I T L X N K A W H O V Z K N S L T X
W S L A B V O K H H E R I N K H S I T W R X A
N R M W H T I O X B N V L T E Z D A V H I S R

D B X O H S B E D W T Z D L W S K A I V H O K
I L V N M Z K W X B E W V S E Z M E O K T R D
M X N S E A T I S W L O A Z M B A X S L R H E

V T M Z I N N B A N K R D T D H M B R Z X K W
H X B E D L K V N X Z I R B S W V S E O I D O
K L V S N A W T B I S O A M D T O R L S H D V

D T A H S O N E X T L E M Z I K T A I S E R K
H B V D Z V K L A R I K H B E O M X B L H Z V
E R N X M S M Z I H L N D X Z K H V T X I B O

D X H S A L O I T L B A K D W T V S O X M N R
H O E B R N V Z B A K V M E T R A I L S X D Z
K D I S L X B O T V W E R H D V Z O B N S O M

X V N N K T R E A Z L K D I M R X L D O W T I
M A S B I O W Z V H E O S B T N W A V Z H M O
K H L N R M W O E I D X A K R E W K S B T E L

Z I X V E W A S B T K O N Z L H B A X M V I Z
I V M X A H S L Z N O K T R D S A W E V X Z I
L E T B S K W E R K A X D I N O W M R N L H K

M H X V A D W N T B S O E H V Z W O I B S A N
I T W O D L X R M I D K L Z A E R T X N H V X
M O S N B R Z V D H R E W V T O B X L S I D X

which, if you can "spot", should help your learning.

The idea, of course, is to reduce your time as much as possible over ten trials."

The cards were always stacked in the same order. The following diagrams (Figure 9) show in what order the cards always appeared, the black numbers indicating the order, the red numbers indicating the cards.

9 1	8 5	7 8	4 2
2 12	6 10	3 11	5 9
10 4	J 7	Q 6	K 3

9 17	8 13	7 14	4 20
2 21	6 23	3 22	5 24
10 19	J 16	Q 15	K 18

9 33	8 29	7 30	4 36
2 25	6 27	3 26	5 28
10 35	J 32	Q 31	K 34

FIGURE 9. DIAGRAM SHOWING THE ORDER IN WHICH THE CARDS WERE ALWAYS STACKED.

(Black figures indicate the order, Red figures indicate the cards).

Answers to the questions and times for each trial were recorded.

TEST VII. Substitution Test B.

The appended sheet of letters, drawn up in the same way as the sheet for Substitution Test A, was given to each subject.

The key was again typed on white paper pasted onto a piece of cardboard 8.2" x 3.5", the key in this case being :

I	L	K	M		V	X		S	E	N	T		H	Z	B	A		W	O	R	D
1	2	3	4		5	6		7	8	9	10		11	12	13	14		15	16	17	18

This key was compiled so that the subject should be able to find it particularly easy to form associations between the letters and the numbers. In

the first group the letters were chosen because each letter had the number of strokes in it indicated by its corresponding number, for example, L has two little strokes in it, therefore its corresponding number is 2. V and X as 5 and 6 were chosen on the basis of sound similarity, and also because of the associations of V as Roman five. SENT as 7, 8, 9, 10 were chosen because each letter was the initial letter of the corresponding number. HZBA as 11, 12, 13, 14 were chosen because of the similarity in shape existing between the letters and the numbers. WORD as 15, 16, 17, 18 were chosen because of the sequence of letters spelling "word", and this being placed at the end of the key, should prove to be an aid to the subject in locating the letters.

The cardboard key was again moved down after each trial, so that the key was always the same distance from the letters during each trial.

The following instructions were read to the subject :

"You have before you a list of letters of the alphabet. In the same way as in the other Substitution Test you are required to write the corresponding number from the key below each letter. In this case the key has been so compiled that certain relationships exist between the numbers and letters which might help you remember them more easily. Thus a certain group of letters might be given numbers corresponding to the number of lines in the letter, e.g., Y = 3, because Y has 3 lines; other groups are selected with other relationships, e.g.

similarity in shape between letters and numbers. Don't devote all your time to looking for such systems; concentrate on learning as quickly as possible, getting as much help as you can from the associations you can discover.

Again 3 lines of letters will constitute a trial. I will give the usual signal "Ready - GO" and you will stop at end of each 3 lines.

Remember to work steadily across the page from left to right."

Ten trials were given, and the time recorded for each trial. The answers to the questions were also recorded.

TEST VIII. Nonsense Syllable Maze (B).

This test, like Test IV, also consists of eleven pages of nonsense syllables, two of which are included here. The material is dealt with in the manner described in the instructions, which are :

"I will give you a sheet of paper on which are 20 lines, each consisting of four nonsense syllables. Again you are required to read them aloud from left to right. When you have reached and said the correct one, I shall say "Yes", after which you will disregard the remainder of that line, proceeding immediately to the next line. On the second and subsequent trials each line will contain the same syllables as before, but arranged in a different order. The correct syllable in a given line is always the same from trial to trial, but it will occur in different positions in the line. You must try

to learn to know the correct syllables for each line. If you remember which syllable in the line is the correct one, say it straight away without first reading the preceding syllables in the same line. If you make a mistake in thus picking out a syllable, try the others in turn as before. The correct syllables in each set of four lines have been chosen for some particular reason. If you notice what the principle is on which a set has been chosen, it should prove easier to remember."

The nonsense syllables were again obtained from Glaze's lists of nonsense syllables, to be found in the Journal of Genetic Psychology, vol. 35, page 265. Again the majority of the syllables were chosen from the 6.67% and .00% "Associative value" columns.

The positions of the correct syllables were arranged according to the same type of plan as that used in Nonsense Syllable Test A. The correct syllables in this test are :

BIP PIB	Chosen because the letters of the second syllable of each pair are those of the first reversed.
CAK KAC	
MEZ MOX MYB MAF	Chosen because each syllable in this group has the same initial letter (M).
ZEG MEF VEC KEX	Chosen because each syllable in this group has the same vowel (E).
WIF JUF TIF CAF	Chosen because each syllable in this group has the same final letter (F).
FYV VEQ QOW WUB	Chosen because the final letter of one syllable is always the initial letter of the next correct syllable.

Again the time taken for each sheet of syllables is recorded and the answers to the questions are recorded.

GUK	BLP	WEZ	JNQ
WOG	ZOV	PIB	ZIS
GAK	KYV	WIX	KEV
VOF	KAG	ZUF	CEB

JIQ	QIF	MEZ	FUV
GEK	YUV	ZIR	MOX
WUG	MYB	BIX	KEZ
GED	CUX	YIB	NAF

KYZ	YIG	ZEG	JYV
QOG	ZIF	GYK	MEF
ZAT	JOQ	YOT	VEG
KIF	TIV	KEX	FUP

VUK	HIF	GID	ZEX
JUF	YIK	FEP	NUK
CIB	VUX	PYD	TEF
ZOP	QUN	ZID	GAF

VUQ	FYV	ZUD	QIK
YOC	ZER	VEQ	YUF
ZIB	QOW	NXP	WUG
YOM	NYB	WUB	JAT

JEQ	WEZ	BIP	GUK
ZIS	TOV	WOG	PIB
KEV	KYV	WIX	GAK
GEB	VOF	KAG	ZUP
FUV	MEZ	JIQ	QIF
MOX	YUV	GEK	ZIR
KEZ	BIX	WUG	MYB
YIB	MAF	GED	CUX
ZEG	KYZ	YIG	JYV
QOC	ZIF	MEF	GYK
ZAT	VEC	JOQ	YOT
FUP	KIF	TIV	XEX
VUK	GID	HIF	ZEX
NUK	FEP	YIX	JUF
FYD	TEF	VUZ	CIB
ZID	ZOP	GAF	QUN
QIM	ZUD	VUQ	FXV
VEQ	YOC	ZER	XUF
MYP	ZIB	QOW	WUC
WUB	JAT	NYB	YOM

SHEET II OF NONSENSE SYLLABLE TEST B.

B. Criticisms of the Tests.

a) Criticisms of the Maze Tests.

While watching the seventy subjects learning the mazes, particularly the linear maze, the writer was often struck by the fact that too many chance factors affected the results obtained to make this a good test of learning ability.

In the record of results of the linear maze are numerous notes such as the following :

(i) Subject got into a blind alley during the seventh trial, into which she had never been before. This threw her out very much; she went backwards practically to the beginning, taking 300 seconds for this trial. From then onwards she thought the general direction of the maze was that shown in Fig. 10 and

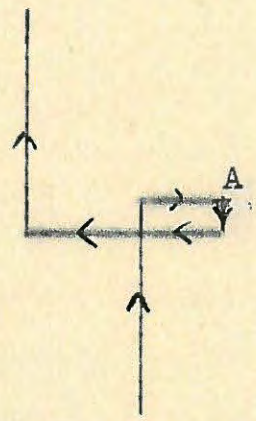

FIG. 10 (See Text).

regularly went backwards at point A.

(ii) Another subject during ninth trial makes a mistake not made previously. This throws S out very much, and puts her into an emotional state.

(iii) Subject makes a mistake, and is thrown out considerably, even though she is using the counting method.

Some subjects used the method of pushing stylus in one direction until stopped, and then changing direction. This method, although it gave the subject little idea of the correct path through the maze, seemed about as effective as any other in reducing time.

The only subjects who really knew the maze by the end of ten trials were those who used the counting method, or said to themselves, "left, left, right", etc. Out of the seventy subjects tested, there were only five who learnt the maze in this way.

A short initial time on the maze often proves to be a disadvantage, not only from the point of view that it is hard to improve upon this time in subsequent trials, but also because the subject is often unaware of many of the blind alleys, and when she gets into them later on, an emotional upset is caused, which may persist throughout the remaining practice periods.

Maze B proved to be a more satisfactory learning task, in that most of the subjects were able to form a fairly good idea of the correct path by the end of ten trials. The excessively long initial and second trials often made it hard to know how to deal with the results. A subject with a long second trial would often show up as a good learner by the formulae we used, even though such a subject did not know the maze very well by the end of ten trials. Compared with this type of reaction is that in which the subject traces the maze quickly the first couple of trials, then realises that she is not really getting to know the correct path, after which she sets about learning the maze, generally resulting in an increase in time.

By the end of ten trials such a subject may have a good idea of the correct path through the maze, and yet the time reduction may have been slight. In one case there was no time reduction after the second trial, and yet this subject knew the correct path through the maze well at the end of ten trials. This also points out one of the disadvantages of using time difference as the criterion for judging improvement (See Chapter VI).

b) Criticism of the card-sorting test.

Another test which the investigator does not feel was satisfactory as a learning task, is the card-sorting test. In this most of the subjects learned the correct pigeon-holes for the cards, and more particularly the order of the cards, much more quickly than was anticipated, and therefore the improvements in time between the second and shortest trials were generally not great. Moreover nearly all the subjects seemed to obtain very similar time scores on this test, and very similar time differences, so that the test was not a good one for dividing the quick learners from the slow. The investigator feels that if a more complicated system were introduced into this test, it could be a very satisfactory learning task.

c) Criticism of the Nonsense Syllable Test.

A possible criticism of the nonsense syllable test is that a subject may take an excessively long time during one of the initial trials, and during that time be repeating the foregoing correct nonsense syllables to herself in order to memorize them. Actually this only happened once on the first nonsense syllable test, and not at all on the second. The

personal influence of the experimenter is probably sufficiently strong to stop subjects from doing this.

d) Criticisms of the Substitution Tests.

In this type of test, the subject sometimes does not seem to memorise the letters and numbers and in this case little improvement in time occurs. However such cases are fairly rare, and those which do occur are probably justly penalised, as surely it is a person with little learning ability who would react in this way.

CHAPTER V.

ADMINISTRATION OF TESTS.

A. Subjects.

As subjects in this research 70 women students of Rhodes University College between the ages of seventeen and twenty-three, were used. These students were all volunteers. Forty-two of them were following the first year course in Psychology, and the majority of other subjects were also undergraduates.

B. When Tests were taken.

Each subject came to be tested individually in a small room 7'6" x 9', which contained two chairs and a table. The room was bright and well ventilated. Each subject was required to come to be tested four times for a 45 minute period. Two tests were administered to each subject during each testing period. A "mental" and a "motor" task were always given during the same testing period, to minimise the effects of boredom and fatigue on the results. The tests were given in the following order :

During Testing Period I	Test I - The Linear Maze.
	Test III - Substitution Test A.
During Testing Period II	Test II - Block-packing Test.
(One week later).	Test IV - Nonsense syllable Maze A.

Forty students were tested during the first two weeks, namely May 5 - 10 and May 12 - 17, 1941 and thirty during the weeks May 19 - 23 and May 26 - 30, 1941.

The remaining four tests were administered

about three and a half months later, in the following manner :

During Testing Period III	- Test V	- Maze B.
	Test VII	- Substitution Test B.
During Testing Period IV	Test VI	- Card-sorting Test.
(One week later).	Test VIII	- Nonsense syllable Maze B.

Again forty students were tested during the first two weeks, namely August 25 - 30 and September 1 - 6, 1941 and the remaining thirty during the weeks September 15 - 19 and September 22 - 26, 1941.

As far as possible subjects came on the same day of the week, and at the same time for successive testings.

C. Instructions.

At the beginning of Testing Period I the following general preparatory remarks were made to each subject :

"This investigation is designed to study the relationships between various kinds of learning. If the results are to be any good, you must apply yourself seriously to learning each of the tasks set you. You must understand that none of these are tests of intelligence, and that we are not interested in comparing your learning ability as a person, with any other student's learning ability. A person who shows a low learning ability in one task may show high learning ability in another; and low learning ability in any task does not indicate low intelligence. Therefore approach these

tasks simply as studies of your progress in learning, in a number of different situations. We are concerned with discovering certain general theoretical relationships, not with discovering the best or poorest learners.

Listen carefully to the instructions in all cases, and try to follow them as closely as possible. In all cases you will be given the signal "Ready - GO"; start as soon as you hear GO.

Please do not discuss or even mention any of the tests to anyone, as this might invalidate someone else's results."

The actual instructions for each test are quoted in full in the previous chapter.

Apart from asking each subject to apply herself seriously to learning each of the tasks, no incentives were used and the times recorded for each trial were always hidden from the subject until the completion of the test.

D. Recording of Results.

For recording results it was found most convenient to use cards 6" x 4" which were set out in the following manner :

NAME OF SUBJECT.	NAME OF TEST.	Questions.
		1a
		b
Age:		2a
		b
1		3a
		b
2		4a
		b
3		5a
		b
4		c
		6a
5		b
		7a
6		b
		c
7		
8		
9		
10		

FIGURE 11. DIAGRAM OF CARD USED FOR RECORDING RESULTS.

In this way ample room was left on the card for calculations and remarks which were not covered by the questions.

The next step to be considered is how to use this data to measure learning ability.

CHAPTER VI.

METHODS OF MEASURING LEARNING ABILITY.

A. Previous Methods of Measuring Learning Ability.

In the previous studies into the inter-correlations between learning abilities, the investigators do not appear to have paid much attention to the way in which they judged such ability. Hall (23) in his study used the total time taken and also the time difference between the first plus second and the thirteenth plus fourteenth trials; while Husband (30) took as his criterion of learning ability the total time taken on the assumption that everyone started with about an equal skill. Both these measures certainly give some idea of the learning taking place, but a very sketchy idea, stressing different aspects. One need only think of a few hypothetical cases in order to see the many factors which both investigators are ignoring.

When time-difference alone is being taken as the measure of learning ability, the way in which that time has been reduced over ten trials, say, is not taken into account. Surely a person who reduces his time greatly during the first couple of trials is a better learner than one who takes nine or ten trials to reduce his time the same amount.

Diagrammatically represented (see Fig. 12), surely A is a better learner than B.

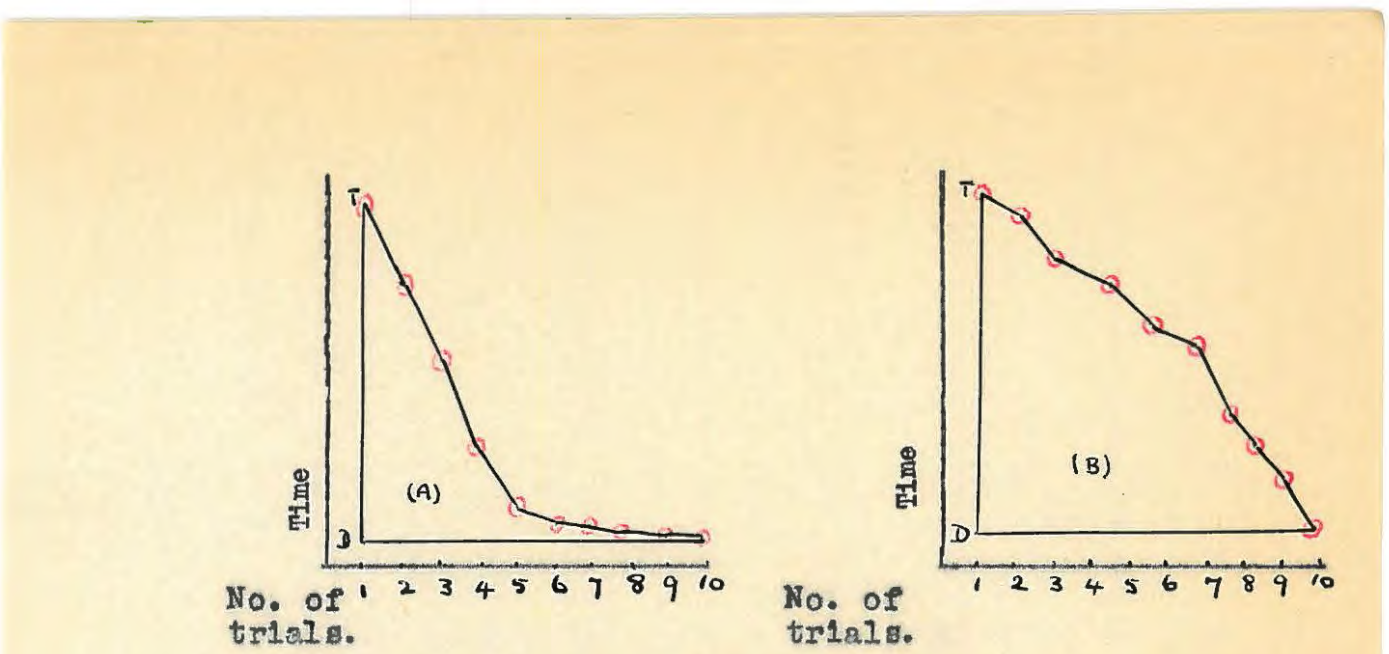

FIGURE 12. (See Text).

By merely taking Time Difference into account, they would both be regarded as having equal learning ability.

Using Time Difference (TD), also favours the person with a poor initial (high time) score, "as the person whose first trial happens to be good is nearer to his ultimate limit, and later increments are achieved only with extreme difficulty." (Husband 30).

Hunter (29 p.554) emphasises the same point when he writes: "As the limits of achievement are reached, and particularly as the performance approximates a physiological limit, the additional increments of accomplishment become more and more difficult to attain."

Husband's method of using the total time taken, overcomes some of these difficulties, but there are also criticisms to be levelled against this method.

In effect what Husband is doing, is taking the total area under the curve, and saying that the smaller this area the better the learning ability.

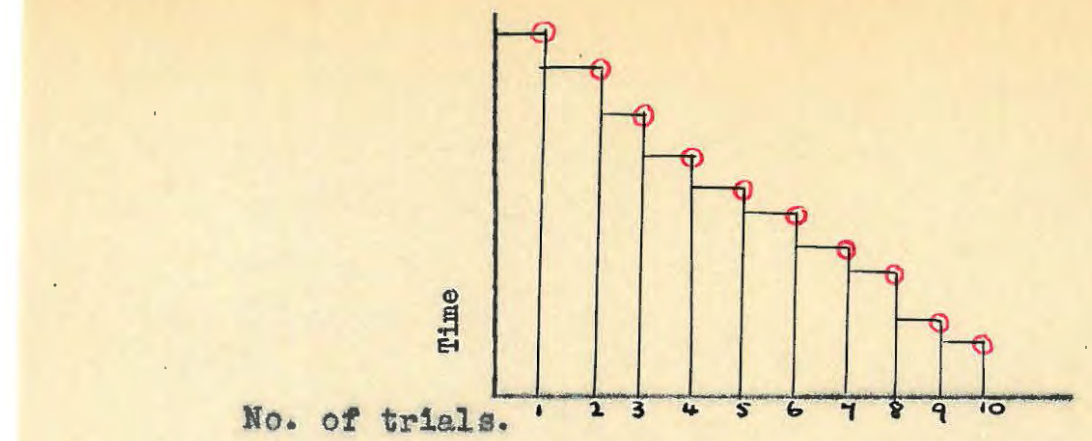

FIGURE 13. DIAGRAM SHOWING THE AREA BEING CONSIDERED BY HUSBAND.

If taking TD favours the person with a poor initial score (high time score), this method certainly does not; in fact, the better the initial score (low time score), the smaller the total area is likely to be. This method of measurement does not seem to measure learning ability solely, as one's initial performance on the task (which has nothing to do with learning ability) affects the results very greatly, as can be seen by regarding the curves shown in Figure 14..

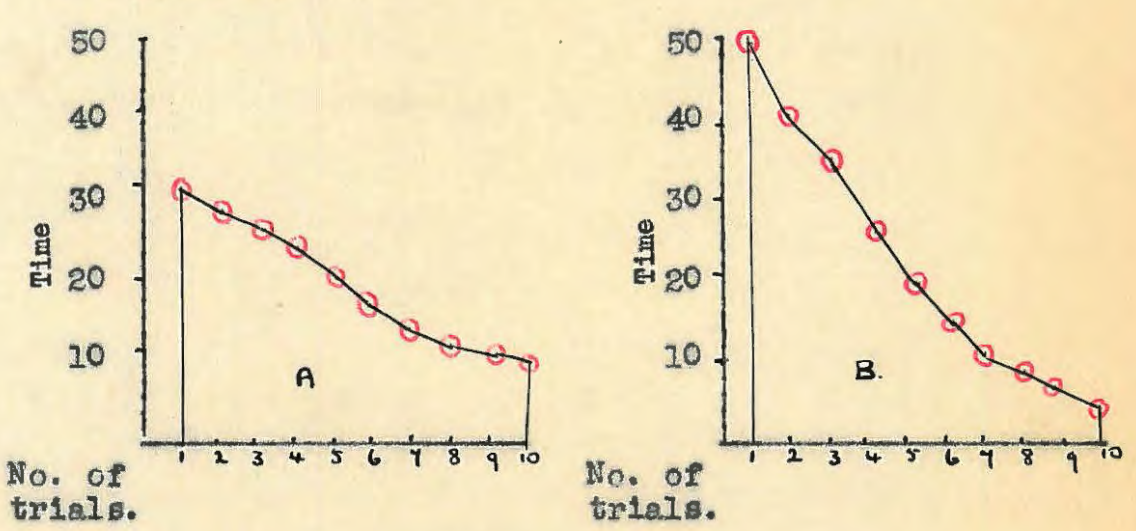

FIGURE 14.

By using as the criterion of learning ability total time taken, A would be the superior learner because the area under the curve is smaller than B's. But far more improvement has taken place in B's case.

He has reduced his time by 45 units, whereas A has merely reduced his by 20. It may be argued that A started with an initially low score and therefore it was harder for him to reduce his time as much as B; but there is a definite rapid improvement taking place in B, which is not in A, and this improvement must surely be regarded as an indication of learning ability.

B. Suggested improvements on these measures.

Because of these considerations the writer spent a great deal of time trying to work out some method which would more satisfactorily measure learning ability. The following are some of the methods devised.

(a) Area under curve between longest and shortest trials.

It was felt that in cases like the above, a truer measure of learning ability could be obtained if the area under the curve between the longest and the shortest trial were taken.

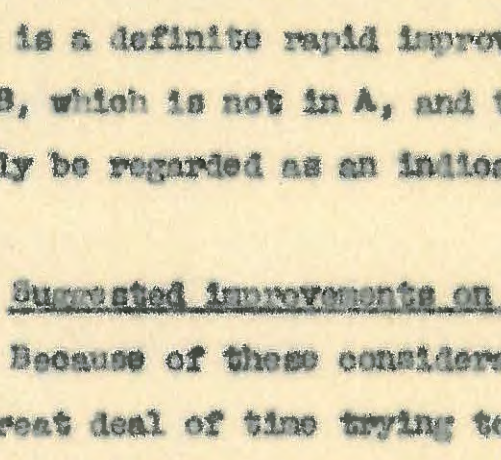

FIGURE 15.

In Figure 15 the distance "a" represents the initial time, and "b" the shortest time. The area under the curve is found in the usual way by adding all

the times taken. The area X is found by multiplying "b" by the number of trials. Subtract X from the total area under the curve, leaving the area A. The smaller the area A, the better the learning ability.

Immediately many objections to this method spring to one's mind. If all the curves started from a common point, and ended at a common point, this method would be applicable; for example, in Figure 16. A would be regarded as a better learner than B and C, and B as a better learner than C.

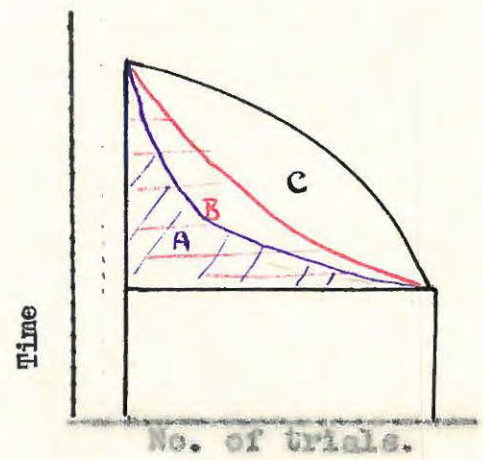

FIGURE 16.

However if the lower point were not common to the three curves, this formula would not give a true indication of the relative learning abilities. Consider the curves shown in Figure 17.

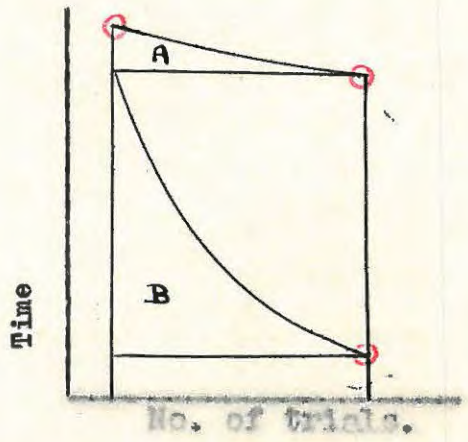

FIGURE 17.

Applying this formula, A would be regarded

as a better learner than B. Yet hardly any improvement has taken place in A and the total time taken is very great. In such cases this measure would be much more inaccurate than Husband's.

The difficulty lies in the fact that no account is being taken of the time reduction. For example (see Figure 18) the smallest possible area obtainable could be represented by a straight line curve, corresponding to "a", the area of which would be zero. Yet in such a case no improvement has taken place at all, but according to the formula such a person would be regarded as a good learner.

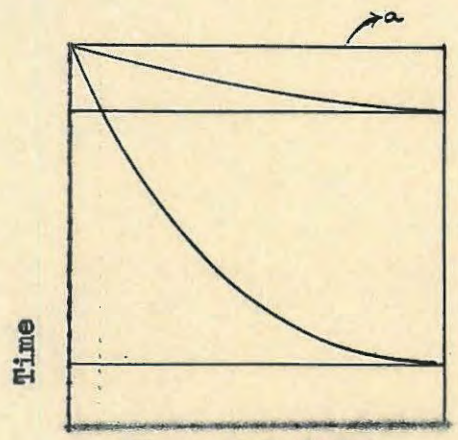

FIGURE 18. DIAGRAM SHOWING HOW A SMALL AREA UNDER THE CURVE CAN BE OBTAINED.

Could a way out of this difficulty be found by combining Time Difference and Area? How about the formula $\frac{TD}{A}$; (A being the area shown in Fig. 15)?

(b) $\frac{TD}{A}$ Formula.

The larger this ratio, the better would the learning ability be, as a large reduction in time, and a small area under the curve indicates that a large time reduction has taken place during a few trials.

There are, however, cases which this formula does not cover, as can be seen in Figure 19.

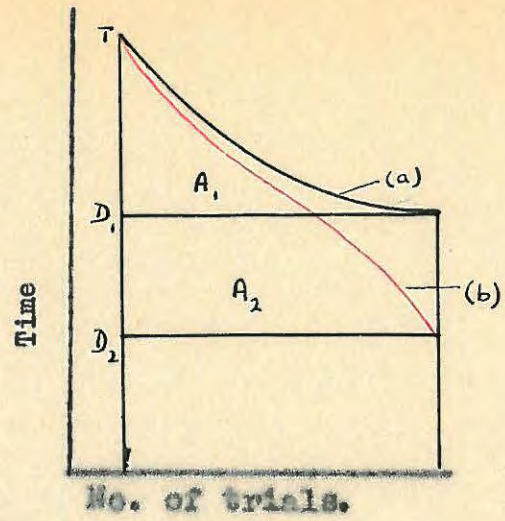

FIGURE 19.

Applying this formula to the curves (a) and (b), $\frac{TD_1}{A_1}$ would be greater than $\frac{TD_2}{A_2}$; and yet (b) has been able to reduce his time more during ten trials, starting from the same initial score than has (a), and therefore his learning ability should be greater. In order to remedy this, the value given to TD should be weighted. $\frac{TD^2}{A}$ may give the desired result.

(c) $\sqrt{\frac{TD^3}{A}}$

After trying many hypothetical curves, it was found that $\sqrt{\frac{TD^3}{A}}$ gave the most reliable results, the hypothetical cases taken not only appearing in their correct order when this formula was applied but the spacing between them being approximately what one would expect, judging learning ability from the shape and position of the curves.

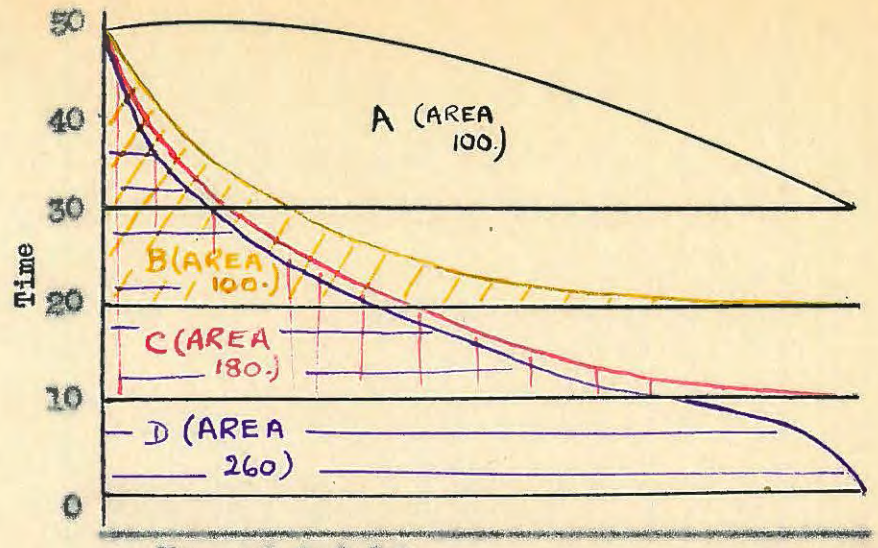

FIGURE 20.

Consider the hypothetical curves shown in Figure 20.

Learning ability measured by TD would be A(20); B(30); C(40); D(50).

" " " " $\frac{TD}{A}$ " " A(·20); B(·30); C(·22); D(·19).

In this case D shows up as the poorest learner, yet he has been able to reduce his time most over ten trials and should be given credit for this. Therefore try weighting time difference, by squaring TD.

Learning ability, measured by $\frac{TD^2}{A}$ would be A(4·0); B(9·0); C(8·9); D(9·6). In this case C shows up as a poorer learner than B, but C has been able to reduce his time more than B over ten trials. Using $\frac{TD^3}{A}$, the results are A(80), B(270), C(355) D(492). Now the four cases seem to occur in their correct order of ability, but B does not, from inspection of the curves, seem to be three times as good a learner as A, and D does not seem to be six times as good a learner as A; that is, the spread has been increased too much by using this formula. By using $\sqrt{\frac{TD^3}{A}}$, the following results are given: A(8·9), B(16·4) C(18·8)

and D(22·0), which would seem a fairly accurate placing and spacing of the respective learning abilities represented by the four curves.

The criticism that can be levelled against this formula is that TD has now been weighted to such an extent that applying this formula to the results makes little difference to the rank order of the subjects, from that obtained by using TD as the criterion of learning ability. The Rank Coefficient was found between the two methods of measurement using the first twenty subjects. Using the substitution test A, ρ was ·98, and on the block-packing test ρ was ·95, which seems to show that this criticism is justifiable.

(d) $\frac{TD^2}{A \times I}$ formula.

Neither TD nor $\sqrt{\frac{TD^3}{A}}$, seems to get over the difficulty quoted by Husband. Both measures seem to favour the person with a poor initial score (long initial time), and the person whose first trial happens to be good, does not have a chance of showing up as a good learner, as he is near to his ultimate limit, and later increments are achieved only with extreme difficulty. It was therefore felt that the initial (or second) score should somehow be introduced into the formula, in such a manner that the person with a short initial time would gain by the introduction, and one with a long initial time would suffer. If initial time (represented as I below, and meaning in each case the time for the second trial on each test as the first trials have been ignored throughout) could be introduced into the denominator of the formula, this

would mean that the smaller initial time would give a larger ratio and therefore a person with a low initial time would have more chance of showing up as a good learner. In this connection the formulae $\frac{TD^2}{A+I}$, $\frac{TD}{A+I}$ and $\frac{TD^2}{A \times I}$ were considered.

Before deciding which of these formulae would be the best to use, eight subjects were chosen from Substitution Test A, on qualitative grounds, and ranked in learning ability, judging this from an investigation of the learning curves obtained and from observations made while they were performing the tests. By this method they were ranked in the following order.

Subjects	Qualitative Estimation	Using TD	Using $\sqrt{\frac{TD^3}{A}}$	$\frac{TD^2}{A+I}$	$\frac{TD}{A+I}$	$\frac{TD^2}{A \times I}$
K	1	1	1	1	3	1
R	2	5	5	5	1	2
P	3	4	2	2	2	3
M	4	2	3	3	5	5
U	5	3	4	4	4	4
F	6	6	6	6	6	6
H	7	7	7	7	7	7
W	8	8	8	8	8	8

TABLE 3. RANK ORDERS OF EIGHT SUBJECTS, USING DIFFERENT FORMULAE.

Using these eight subjects $\frac{TD^2}{A \times I}$ ranks them more closely to the qualitative estimation than do the formulae $\frac{TD^2}{A+I}$ and $\frac{TD}{A+I}$.

The Rank Coefficient was found between the $\frac{TD^2}{A \times I}$ and TD methods of measurement, using the first twenty subjects on the substitution test A. On this

test ρ was ·93 and on the block-packing test ρ was ·73. ρ was also calculated between $\frac{TD^2}{A \times I}$ and $\sqrt{\frac{TD^3}{A}}$, being ·96 on the substitution test, and ·87 on the block-packing test. In both cases the Rank Coefficient is highest between TD and $\sqrt{\frac{TD^3}{A}}$, being in one case ·98 and in the other·95, while it is lowest between TD and $\frac{TD^2}{A \times I}$, being ·93 and ·73.

Can we conclude from this that $\sqrt{\frac{TD^3}{A}}$ does not differ greatly from plain TD as a measure of learning ability and that ranking subjects by this method will give virtually the same results as using Time Difference? The fact that $\frac{TD^2}{A \times I}$ does not correlate so highly with either of these measures may mean that it is taking account of more factors, such as the shape and position of the learning curves obtained, and is therefore a better measure of learning ability.

To show the improvement of the $\frac{TD^2}{A \times I}$ formula over TD, consider the following two sets of results which were obtained on the Maze B test.

	A	B
1.	34.8	113.7
2.	20.8	221.9
3.	24.3	190.8
4.	23.0	105.1
5.	21.7	39.5
6.	18.0	64.4
7.	16.3	48.0
8.	16.1	91.7
9.	15.5	32.0

TABLE 4. RESULTS OF TWO SUBJECTS ON MAZE B.

(The first scores in each case have been omitted.)

A has a TD of 19.3 seconds, while B has a TD of 81.7 seconds. From this criterion it would seem as if B were a better learner than A; yet from observations made by the investigator while watching the two subjects doing the test, it was obvious that A very soon formed an idea of the correct path through the maze, while B had little idea of the true path, even after ten trials. If the $\frac{TD^2}{A \times I}$ formula is applied to these two sets of results, $\frac{TD^2}{A \times I}$ for A comes to ·2099, while for B it is ·0948. This seems to be a more just reflection of the relative learning abilities of the two subjects than the results obtained by using TD.

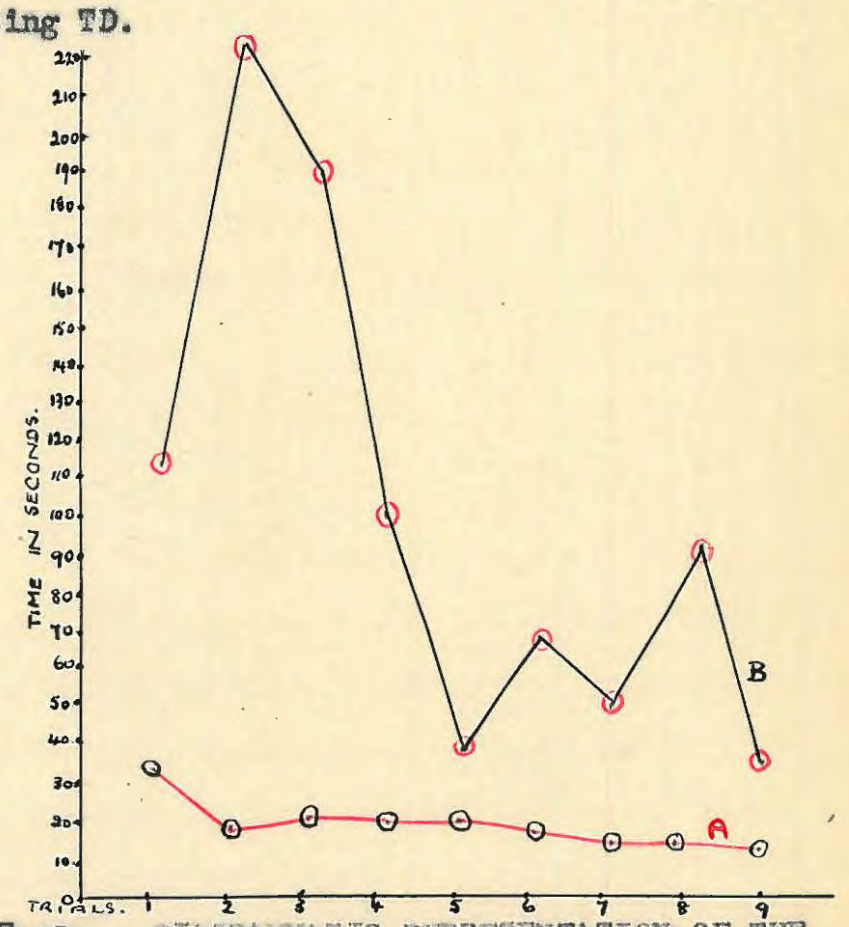

FIGURE 21. DIAGRAMATIC REPRESENTATION OF THE LEARNING CURVES OBTAINED BY A AND B, ON MAZE B.

Are the correlation coefficients low due to the fact that the TD method is an unsatisfactory means of measuring learning ability? The present investigator

thought that this might partly account for the small size of the correlations obtained as very often someone would get a large TD on one test due to the fact that the time taken for the first couple of trials is so long that without really getting to know the task much better, she would be able to reduce her time considerably. This type of behaviour occurred frequently in the maze tests. On another test such people's reactions were generally rather different. For example, poor learners on the substitution test generally showed little reduction in time, and therefore obtained a small TD. In this connection consider the results of the following two subjects, C and D.

	Maze B (TD in secs.)		Substitution B (TD in secs.)	
C.	216	Range from 0 to 256; mean 47.5.	29	Range from 12 to 109, mean 45.4.
D.	253		32	

A few such subjects would immediately lower the correlation between the two tests and quite unjustifiably, as such people in reality possess poor learning ability on both tasks.

The $\frac{TD^2}{A \times I}$ formula does not seem very suitable for such extreme cases as these, for A X I in such cases is not large enough to compensate for the extremely large proportions which TD^2 assumes. For example, taking the two results just quoted, and applying the $\frac{TD^2}{A \times I} \times 1000$ formula, they become

	Maze B ($\frac{TD^2}{A \times I} \times 1000$)		Substitution B ($\frac{TD^2}{A \times I} \times 1000$)	
C.	603	Range from 0 to 634.	34	Range from 22 to 203.
D.	471	Mean 171.	70	Mean 91.

This shows that applying this formula has not improved

the position materially. After considering cases such as this, it is perhaps not surprising that the correlations found using the $\frac{TD^2}{A \times I}$ formula do not differ much from those found when using TD as the criterion of learning ability. (See Chapter VII).

It also occurred to the writer that working from a smoothed curve might be an advantage, but when considering curves such as A and B, quoted previously, it is doubtful whether this would make much difference. Dealing with curve B, the difference between the average of the first three trials and the last three trials is 118·2, and on curve A this difference is 10·7. Here the position has been aggravated, not improved by using smoothed curves. In A's case, where the improvement has been steady at the beginning, this smoothing of the curve has penalised the subject for the initial time is made lower and therefore TD is reduced from 19·3 to 10·7. In B's case, where the time has increased during the first three trials, the initial time is increased and TD is increased from 81·7 to 118·2. The only cases in which smoothing the curve would be of some advantage, are those where only the second (regarded as initial) time has been long, due to chance factors. The disadvantages outweigh the possible advantages of working from smoothed curves, and this technique was therefore not adopted.

(e) $\frac{TD}{\text{Total Area}}$ formula.

As A X I in the formula $\frac{TD^2}{A \times I}$ did not prove to be large enough to compensate for the proportions which TD^2 assumed, when TD happened to be large, due to an extremely long initial time, perhaps using the

total area under the curve (in the manner Husband did), and combining this in some way with Time Difference, will get over this difficulty. $\frac{TD}{\text{Total area}}$ was tried on the results A and B quoted previously. A becomes $\frac{19.3}{190.5} = \cdot10$, and B becomes $\frac{81.7}{907.1} = \cdot09$. This seems satisfactory in that the order of the two subjects is reversed. Even though this formula may in practice give one a fairly good conception of people's relative learning abilities, there are grave criticisms which can be levelled against it on theoretical grounds. Consider two curves, in shape exactly similar, the one (X) commencing and ending with a larger time score than the other (Y), as shown in Figure 22.

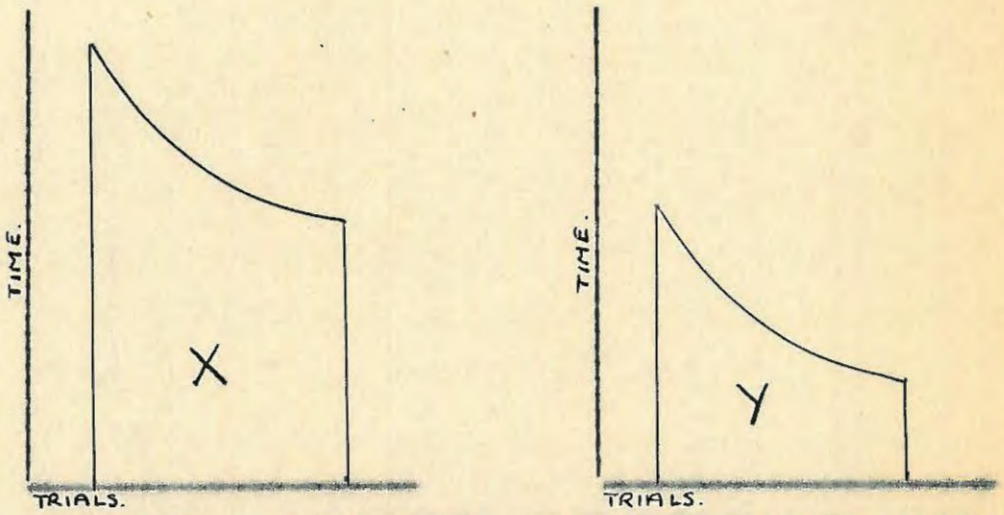

FIGURE 22. TWO CURVES, EXACTLY SIMILAR IN SHAPE.

Should X be penalised because he has less performance ability on the task? The shape of the two curves is exactly the same, and the time difference is the same. The argument in favour of penalising X is that X is working nearer to his ultimate limit, and the time reduction is being achieved with more difficulty than X's. But what evidence have we that this is so? X's physiological limit may not be the same as Y's, and it is only if we assume that everyone has the same

physiological limit that we can apply this formula. This criticism seemed such a grave one that this formula was not used.

C. The Common Points of Mastery Technique.

In an article by Floyd L. Ruch (50) entitled "The method of common points of mastery as a technique in human learning experimentation", he points out that a comparison of groups or individuals with regard to learning ability is frequently rendered peculiarly difficult because of the fact that initial ability varies between the groups or individuals. For example, one subject starts at a level of ability indicated by a score of 40 units of work per trial and improves over the period of practice to a score of 60 units. Another subject starts at 30 units and improves in the same number of trials to 50. Both subjects have an absolute gain of 20 units. If we express improvement as a percentage of the initial score (which is one traditional method of finding learning ability), we must give the first subject a rating of 50%, and the second a rating of 66-2/3%. If we employ the Time Difference or absolute gain method, we must conclude that the subjects are equal. Both methods are invalid. The percentage method is so because it assumes that the arbitrary zero point of the test coincides exactly with the absolute zero point of the ability being considered. The method of defining learning ability as the difference between the initial and final performance of the subject is rendered invalid, as we have already shown, by the fact that units at different levels of performance probably have different values. A raw score difference of 10 points at one end of the scale is not necessarily equal to a

raw score difference of 10 points at the other end of the scale. The elimination of the first error in a learning experiment does not have the same significance as a measure of ability as the elimination of the last error.

Ruch holds that the method of common points of mastery will surmount these difficulties. There are two methods which he advocates. Preliminary to using either method the learning curves based on the trial scores for each subject are smoothed by a running average covering enough points to eliminate serious chance fluctuations in the scores. Then a particular level of performance common to the early trials of all subjects is determined by inspection.

In Method A a second or final common point of mastery is selected. This point must also be common to the curves of all subjects. Learning ability is then defined as the number of trials or number of errors or amount of time required to improve from the initial to the final common point of mastery. "The only assumption that need be made here", says Ruch, "is that ten trials represent more effort than nine trials."

In Method B the initial common point of mastery is established in the same way as before and learning ability is defined as the amount of improvement in the first segment of trials beyond this common point.

Preliminary investigations indicate that the method of common points of mastery has considerable promise in the study of the intercorrelations between the abilities to learn different tasks. Applying this method to the results obtained by Hollingworth

(27) for his sixteen subjects who learned three-hole motor co-ordination and opposites, the intercorrelation using the common points technique was found to be +.59, while using the methods "Difference between first and fourth blocks of five trials", "First and second blocks of five trials", and "Third and fourth blocks of five trials", the correlation coefficients were respectively -.14, -.17 and -.11.

Ruch accounts for this increase in the following way: "To make the explanation simple, let us assume that there is a correlation between innate .... ability to learn two different tasks. In the typical learning experiment the subjects are far from unpractised at the moment the conditions of a formal learning experiment are imposed upon them.... What we know definitely is that subjects show individual differences in performance on the very first trial of the formal experiment. One subject might have a high initial status because he has had much unrecorded prior practice though he may be only average in native learning capacity. Another subject might make a low initial score because he is average in native capacity and has had but a small amount of prior practice. A third subject may stand high on the first trial by virtue of great native capacity even though he has had little unrecorded prior practice. These are just a few of the combinations of prior practice, native capacity and initial performance. Add to this motley of variable unknowns the further fact that absolute gains have different meanings in terms of ability.at different ranges of the scale, and it would appear that in the traditional method we are dealing with enough attenuating and distorting factors to conceal a fairly

sizeable correlation. The method of common points eliminates variable initial ability and allows gain to be determined systematically by native capacity beyond that point." (Rush 50).

There are, however, some criticisms which can be levelled against this technique, particularly in its application to material of the kind being dealt with in this investigation. In simple learning tasks such as were used, it would be hard to fix a common point so that the people who passed through it would all still be showing fairly rapid improvement. For example let us again take two curves of exactly the same shape and time reduction but at different levels, as shown in Figure 23.

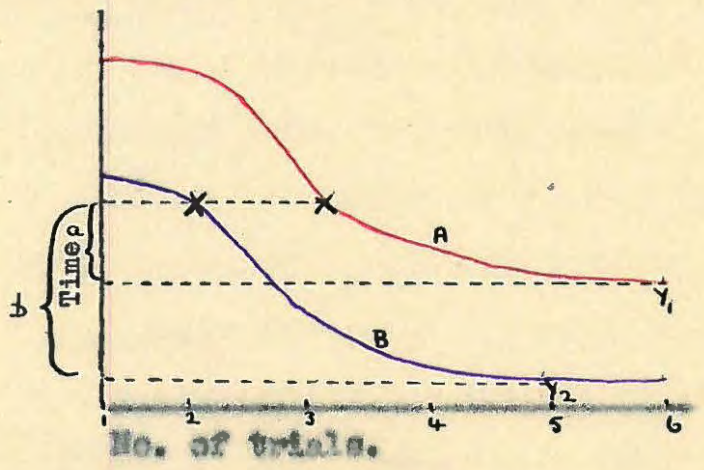

FIGURE 23.

A and B both pass through the common point of mastery X, A on his third trial, and B on his second. Taking learning ability as the improvement over the next three trials, A's learning ability can be represented by the distance "a" and B's by "b". By this criterion A has less learning ability than B. A's performance ability may be poor, but he has learnt just as quickly as B. This does not show up, however, because the common point chosen was at a "slow" part of his curve.

In work requiring prolonged practice and in

which improvement is steady and gradual, this method would probably be satisfactory, but it could not be used in this investigation as the subjects were only given ten trials on each of the various tasks, and the results of numerous subjects would have had to be eliminated, due to the fact that no common point of mastery was reached. In this connection consider Table A in Appendix I, where it can be seen that the highest shortest-times are in every test greater than the lowest initial (second) times, which means that no common point of mastery could be obtained from these curves.

A slight improvement on Ruch's method A might be suggested. Consider the curves shown in Figure 16.[p89] These three curves all pass through the common points of mastery X and Y, and each subject takes four trials from the initial point to reach the final. If only the number of trials were being considered as the criterion of learning ability, all three subjects would be considered to have equal ability. A, however, has reduced his time more quickly than B and C, and B more quickly than C; therefore area under the curve, or total time taken between the two points should also be taken into consideration. Perhaps this is what Ruch means when he writes: "Learning ability is then defined as the number of trials or number of errors or <u>amount of time</u> required to improve from the initial to the final common point of mastery", but the point is not elaborated.

D. <u>Measuring Learning Ability by Retentivity.</u>

It is perhaps not necessary to define learning ability as progress, or <u>rate</u> of learning. Thorndike

(50 page 5) thinks of learning as acquiring responses and changing the strength of their connections with the situations of life. Therefore learning ability could perhaps be regarded in terms of the greatest attainable efficiency of the connections, and not as the rate at which connections are formed. As a corrollary of this, learning ability could perhaps be defined in a negative way in terms of the rate of loss of efficiency, after learning to a certain stage has taken place. It will be argued that this measures merely retentivity. But learning ability is not solely the speed of educing relations - that is what is tested in an intelligence test - it is surely the ability to educe relationships relevant to some goal, and to <u>retain</u> them so that they follow rapidly in a certain sequence.

If one is prepared to accept this as a definition of learning, why not train all subjects to a certain standard of efficiency (i.e. to a final common point of mastery, but not beyond), then let a definite period of time, say a week, elapse, after which the subject is required to do a trial, or a few trials on the same task, to see how much he has retained during the interval.

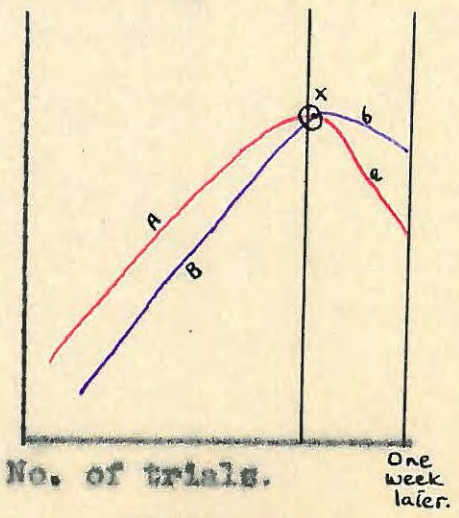

FIGURE 24.

In Figure 24 A and B are both trained to reach the standard of efficiency indicated by the common point of mastery X. After a week's interval they both do another trial on the task, their respective powers of retentivity being represented by the lines "a" and "b", A being classed as the poorer learner because he has lost in efficiency more than B; that is, his powers of retentivity, on that particular test, are less.

This method was not used in this investigation, due to practical difficulties in the way of training all subjects to a common point of mastery, in a limited length of time. It was often difficult to fit two tests into a forty-five minute period, giving only ten trials on each task, and the unpredictability of the length of time it would take to train a given subject to the final common point of mastery, made this suggestion impracticable for use in this investigation. However the writer has outlined this technique here as she thinks some use of the method may open up interesting avenues of research into a slightly different aspect of learning ability.

In the following chapter the results obtained from the application of certain of the formulae discussed here, will be given.

CHAPTER VII.

TREATMENT OF DATA, AND RESULTS OBTAINED.

In the previous chapter various methods of measuring learning ability were discussed. On theoretical grounds the $\frac{TD^2}{A \times I}$ formula seemed as if it would be most satisfactory. This formula, as well as the TD, has been applied to the data, to find out which formula would give the higher correlation coefficient.

A. TD and $\frac{TD^2}{A \times I}$ Formulae applied

Firstly the TD's were calculated for the 70 subjects, on each of the eight tests. TD in each case being taken as the difference in time between the second trial, and the shortest trial.

$\frac{TD^2}{A \times I}$ was then calculated for everyone on each of the tests. TD^2 is self-explanatory. A is found by adding the times taken for trials 2 to 10 and subtracting from this the shortest time multiplied by 9 (i.e. the number of trials). Diagrammatically represented it is the area shown in Fig.25; shaded area being A, the 6th trial being the shortest.

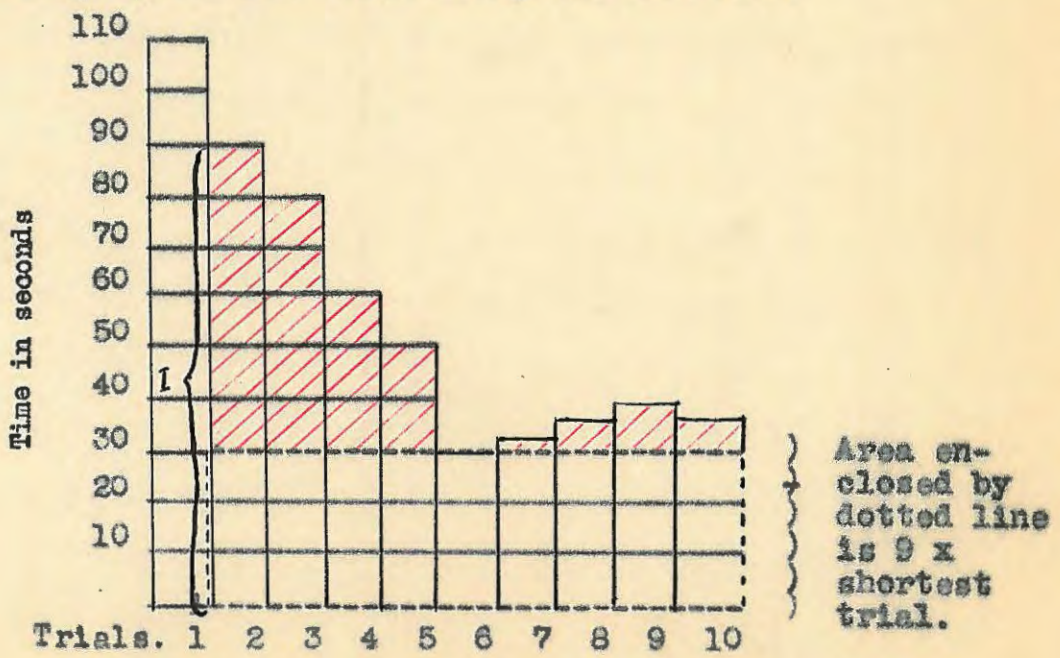

FIG.25. SHOWING AREA CONSIDERED AS A.

I is always the time taken for the second trial as the first trials on learning tasks are notoriously unreliable (Hall, 23). For convenience sake these results were all multiplied by 1000, to avoid working with decimals.

B. The full table of correlations, and distribution of coefficients.

Using TD as the criterion of learning ability, the following is the full table of correlations among the learning tests. The method used for calculating "r" is that outlined by Holzinger (28) and shown in Appendix III.

	Maze A.	Sub.A.	Blocks.	Non A.	Maze B.	Sub.B.	Cards.
Sub.A.	+.00						
Blocks	+15	+.08					
Non A.	+.08	-.04	-.15				
Maze B.	+03	+.09	-.02	-.07			
Sub.B.	+08	+.26	-.09	+.14	+.25		
Cards B	-.20	+.26	+05	-.06	+05	+18	
Non.B.	-.01	-.07	-.11	+40	+02	+34	+.13

TABLE 5. INTERCORRELATIONS BETWEEN ALL THE LEARNING TASKS.

The size of these correlation coefficients is remarkably similar to the size of those obtained by Husband (30). The median coefficient found by Husband was +.13, which led him to the conclusion that learning was not a single function, but rather existed in the plural. The median of the correlations tabulated above is +.05 (Mean .06), which seems surprisingly low, considering that the tests were longer than those used by Husband, and also considering the great similarity of the tests in the Facilitation and Achievement groups.

TABLE 6. DISTRIBUTION OF COEFFICIENTS.

Correlation range.	No.of r's found from this investigation.	No.of r's found in Husband's investigation.
·50 to ·59	0	1
·40 to ·49	1	1
·30 to ·39	1	6
·20 to ·29	3	26
·10 to ·19	4	19
·00 to ·09	9	23
-·10 to -·01	7	14
-·20 to -·11	3	1

The distribution of coefficients found from this investigation coincides almost exactly with that quoted by Husband, except that the majority of these correlations are even lower than his. The few slightly negative coefficients can be considered inconclusive, as can those which are slightly positive. Those which fall within the range of -·10 to +·10, and even up to $\pm$ -·20, are so near their P.E's that the trends prove little. Four-fifths of our correlation coefficients lie within the range $\pm$·20.

Taking the criterion that an r to be significant, should be at least three times its P.E, it can be seen that only five figures attain the required magnitude.

C. Intercorrelations among Motor, Mental, and Ideational Tests.

Husband divides his correlations into three groups. The median between the motor tests is +19, among the rote learning tests is +·20, and it rises to

+·25 with the three correlations between ideational tests.

The results of this investigation have also been divided into those between motor, "mental" and ideational tests, so that they may be compared with Husband's three groups. The only correlation coefficient that is comparable with the median of those between Husband's ideational tests, is that between Nonsense Syllables B, and the substitution test B, as these are both mental tasks, involving relation eduction.

TABLE 7. INTERCORRELATIONS AMONG MOTOR TESTS.

	Maze A.	Blocks.	Maze B.	
Blocks	+·15			
Maze B	+·03	-·02		
Cards	-·20	+·05	+·05	Median +·04

TABLE 8. INTERCORRELATIONS AMONG "MENTAL" TESTS.

	Sub.A	Non.A.	Sub.B.	
Non.A.	-·04			
Sub.B.	+·26	+·14		
Non.B.	-·07	+·40	+·34	Median ·20

TABLE 9. INTERCORRELATION BETWEEN IDEATIONAL TESTS.

	Sub.B
Non.B.	+·34

The median correlation between the motor tasks is much lower than that obtained by Husband, i.e. +·04, as compared with his +19. The median of the mental tasks is +·20, which is the same as Husband's median between his rote learning tasks. The correlation obtained between substitution test B, and nonsense syllables B, is +34, which can be compared with his median of +·25 among ideational tests. Thus

it would seem that correlations are slightly higher among the learning tasks involving more complex functions, than among simple motor tasks.

D. Intercorrelations Among Facilitation, And Achievement Tests.

The correlation coefficients obtained for the Facilitation tests, using both the TD, and $\frac{TD^2}{A \times I} \times 1000$ criteria of learning ability, are shown in Table 10.

TABLE 10. INTERCORRELATIONS BETWEEN THE FACILITATION TASKS.

Using TD			Using $\frac{TD^2}{A \times I} \times 1000$		
	r	P.E		r	P.E
M vs. S	-.003	± .08	M vs. S	+.03	±.08
M vs. B	+.15	± .08	M vs. B	+.16	±.08
M vs. N	+.08	± .08	M vs. N	-.09	±.08
S vs. B	+.08	± .08	S vs. B	+.02	±.08
S vs. N	-.04	± .08	S vs. N	+.08	±.08
B vs. N	-.15	± .08	B vs. N	+.11	±.08
Average	+.02		Average	+.05	
Median	+.04		Median	+.055	

The highest correlation obtained is +.16, which is only twice its P.E., therefore none of the correlation coefficients found between the Facilitation tests, is significantly different from zero. Neither are any of the correlations significantly changed by using the $\frac{TD^2}{A \times I}$ formula, instead of TD. The mean correlation using TD is +.02, and using $\frac{TD^2}{A \times I}$ is +.05.

Correlations computed in the same manner, using the Achievement tests, are shown in Table 11.

TABLE 11. INTERCORRELATIONS BETWEEN THE ACHIEVEMENT TASKS.

Using TD			Using $\frac{TD^2}{A \times I}$ x 1000		
	r	P.E.		r	P.E.
M vs.S	+·25	±·08	M vs.S.	-·02	±·08
M vs.C	+·05	±·08	M vs.C	+·12	±·08
M vs.N	+·02	±·08	M vs.N	-·13	±·08
S vs.C	+·18	±·08	S vs.C	-·08	±·08
S vs.N	+·34	±·07	S vs.N	+·06	±·08
C vs.N	+·13	±·08	C vs.N	-·18	±·08
Average	+·16		Average	-·04	
Median	·165		Median	-·05	

The highest correlation obtained is +·34, which is nearly five times its P.E., and therefore significantly different from zero; while the r of ·25 is three times its P.E. The average correlation using TD is +·16, while using $\frac{TD^2}{A \times I}$ it is -·04.

Using the TD criterion seems to bear out the statement made previously, that correlations are slightly higher between the learning tasks involving more complex functions,(achievement tasks) than between simple facilitation tasks.

However using the $\frac{TD^2}{A \times I}$ criterion, this does not seem to be borne out. In both cases the difference is so small, and the median correlations themselves differ insignificantly from zero, that the small trends cannot be regarded as meaning much.

It was pointed out in the previous chapter how the $\frac{TD^2}{A \times I}$ formula could be a more inaccurate criterion of judging learning ability than plain TD, and after the following points had been considered,

the $\frac{TD^2}{A \times I}$ formula was no longer used.

E. Comparison of results obtained by using the TD and $\frac{TD^2}{A \times I}$ formulae.

Correlations were calculated between TD and $\frac{TD^2}{A \times I}$, using the same tests, and all 70 subjects. The following r's were found :

TABLE 12. SHOWING CORRELATIONS BETWEEN THE TD AND $\frac{TD^2}{A \times I}$ METHODS OF MEASUREMENT.

		r	P.E.
Using	Substitution Test A.	+·81	±·03
	Nonsense syllables A.	+·76	±·03
	Block Packing.	+·77	±·03
	Maze A	+·70	±·04

These correlations are all high, therefore it would seem that it does not matter very much whether one uses the TD formula or the $\frac{TD^2}{A \times I}$ formula as the criterion for judging learning ability. On inspection it would seem that the $\frac{TD^2}{A \times I}$ formula affected the results of the Maze Test more than any of the others, in that the correlation between $\frac{TD^2}{A \times I}$ and TD on this test was the lowest. The significance of the difference between the two correlations found on the substitution test (the highest correlation) and on the maze test (the lowest correlation), was calculated in the manner outlined by Shepherd Dawson (14 pg.138). $\frac{X}{PE}$ in this case was 2·5, and not 3, which it must be to reach the 5 per cent level of significance.

Therefore, as the correlations between the two methods are all high, and as the $\frac{TD^2}{A \times I}$ formula

does not affect the correlation coefficient in a significantly different manner when applied to different tests, it was decided from this stage onwards, only to use TD as the criterion for judging learning ability, as it involved less calculation, and would indicate as well as the other formula any clear trends.

F. Correlations between the corresponding tests, in the Facilitation and Achievement groups.

TABLE 13. CORRELATIONS BETWEEN CORRESPONDING TESTS.

	r.	P.E.
Sub.A vs. Sub.B.	+.26	±.08
Non A vs. Non B.	+.40	±.07
Maze A vs. Maze B.	+.03	±.08
Blocks vs. Cards.	+.05	±.08

Even these correlations are surprisingly low, considering the great similarity of the operations involved in both substitution tests, both nonsense syllable tests, and both mazes. The correlations between the two nonsense syllable tests, and between the two substitution tests, are signigicantly different from zero.

The extremely low r between the two mazes is of particular interest, in view of the statement made by Hall (23 pg.185) "It would appear as though a rather important general maze learning ability did exist". The lowness of the r may in part be accounted for by the fact that Maze A seemed to be rather an unsatisfactory learning task (See Chapter IV), to learn in ten trials. However mazes of this type

are fairly well standardized, and the correlation which is virtually zero, between the two stylus mazes, must be accounted for by the fact that the opportunity to educe relations afforded by Maze B, affected the results so materially, that no correlation, due to the similarity in manual operation involved, could show up; or because Maze A in particular, or mazes in general are unreliable as tests of human learning ability; or because learning abilities are specific to the material being learnt.

Another statement made by Hall comes to mind when considering these low correlations - namely, "one could make out a very good case for specialized learning abilities if low correlations were obtained between very similar learning tasks". In our case the correlations between similar motor tasks are very low, but between the mental tasks, although certainly not in the neighbourhood of 1, the correlations are significantly different from zero. This again seems to give evidence in support of the statement that correlations are higher among learning tasks involving more complex functions, than among simple motor tasks. The highest correlation is between the two nonsense syllable tests, which involve the least motor activity. The other fairly high correlation exists between the two substitution tests, in which the learning which occurs is probably mainly "mental", but which involves a fair amount of motor response, in the form of writing down the requisite numbers.

The low correlations between the motor tasks may be due more to differences in work method than to differences in learning ability. In an article by van Dusen (16 pg.225), the significance of work methods in

learning the motor skill of card-sorting, is made clear.

At the beginning of this investigation (see Chapter II) it was decided to try and choose eight tests, four of which, as far as possible, involved pure facilitation prominently, in the learning process required, and the other four involving achievement prominently. If the two groups of tests could be kept very similar, apart from the amount of achievement that could be used in one set (achievement always meaning ability to educe relations, reasoning, etc.) this would make it easier to interpret the correlations found.

If a fairly high correlation were found between the various problems devised as facilitation tasks, and also a high correlation between achievement learning tasks, but not a very high correlation between the two sets of tests, then the reason for this correlation would be the difference between the processes required to make a good facilitation, and a good achievement learner. Looking at the correlation coefficients obtained between the various facilitation tasks, it seems as if there is virtually no correlation at all. Three of the coefficients are slightly positive, and three are slightly negative, the median coefficient being + ·04, with a P.E. of $\pm$·08.

The achievement tasks correlate a bit more highly, all the coefficients being positive, and the median being + ·17. One of the correlations is five times its P.E., and another three times.

From this it might be inferred that the abilities needed to make an achievement learner, are

more generalised, than the abilities used in learning the facilitation tasks. It would be interesting to know how the learning abilities on the various tasks correlate with intelligence. Would the abilities on the achievement tasks correlate more highly than those on the facilitatiom tasks?

The theoretical considerations of the correlations found between the similar tests in the facilitation and achievement groups, have been discussed and it was pointed out how extremely hard it was to know whether the low correlations obtained between the similar motor tasks, were due to the fact that the B tests gave an opportunity to educe relations. The interpretation of these low correlations is made even more difficult by the fact that the block-packing, and card-sorting tests differed more in the type of operation involved, than did the other three pairs of similar tests.

In the two pairs of tests which do show a fairly high correlation: is the size of this correlation, and the fact that it does not approach 1, due to differences in amount of facilitation and achievement used; or due to the fact that learning ability is specific to the particular task being learnt; or perhaps due to differences in emotional attitude? In some subjects there may have been a positive transfer, and doing a test of a similar nature has aided their learning of the second test. In others there is a negative transfer. Several subjects told me they could still remember the numbers and letters of the first substitution test, although several months had elapsed before the second one was given. Many

subjects who had evoked interest on the first test, were bored by having to do another test of a similar nature. From questions asked the subjects, as to their work methods on the second batch of tests, the large majority of people, did, at least in part, use the "achievement" method of learning. Which of these factors is the important one, or whether they are all important in keeping the correlation coefficient fairly low, on two very similar tests, it is hard to say. A fruitful line of research would be an attempt to isolate such factors, and this could be done more easily by using very similar tasks. Previous investigations using very different learning tasks, have not been successful in demonstrating the existence of a general learning ability. When dealing with such a question it is the obvious thing to try to find a correlation between tasks that are as different as possible, but since this has not been found, it is time that people tried working from the inside outwards, in an attempt to isolate disturbing influences in the learning situation.

Summary.

1. 28 correlation coefficients between eight tests of learning were calculated. The median coefficient was +.05.

2. The intercorrelations were all very low, only 5 out of 28 being significantly different from zero, and four-fifths of the r's being so near their P.E's that the trends prove little.

3. Considering the correlations obtained between motor, "mental" and "ideational" tasks, it would seem

that correlations are slightly higher among learning tasks involving more complex functions, than among simpler motor tasks.

4. This statement also seems to be borne out by the correlations found between corresponding tests in the "facilitation" and "achievement" groups. The r's found between corresponding mental tasks were significantly different from zero, while the motor tasks virtually showed no correlation at all.

5. Evidence in support of this statement is also found by comparing the correlations obtained between the "facilitation" tasks, and the "achievement" tasks, the median for the latter being +·17, and therefore higher than for the former, which is +·04. These results are very inconclusive, but it might be argued from them that the processes needed to make an achievement learner are more generalised than those needed to make a facilitation learner.

6. The results obtained in this investigation are very similar to those obtained by Hall (23) and Husband (30), the correlation coefficients in this case being even lower than those reported by them.

G. Are the low correlations due to Emotional Factors?

Is the lowness of these correlations due to the fact that some subjects react peculiarly to the testing situation? In other words, are the low correlations perhaps due to the fact that emotional factors are causing the subjects to react in peculiar ways to the testing situation, rather than due to differences in learning ability? This point occurred to the writer as a result of observations made while

testing the subjects on the set of facilitation tasks.

In order to try and get some qualitative evidence on this problem, the lists of questions, quoted in Chapter IV (see p 66) were asked the subjects, when they came to be tested on the achievement tasks. By this method it could be seen who reacted in an extraordinary emotional manner to a given test, and people who had been affected in this way could be eliminated, and correlations calculated only using the remaining subjects. This elimination must be done solely on the basis of answers to the questions, and observations of the experimenter, recorded while the subject is performing. It would be quite unjustifiable to eliminate subjects by an examination of their scores.

To do this the following technique was applied. Using the substitution test B, and the Nonsense Syllable test B, all people were eliminated who either actively disliked the test or liked it very much indeed; who disliked tests involving figures, or disliked reading aloud; who were over-cautious and nervous; and who found that doing the previous test first was a disturbing factor. Regarding the list of questions asked, this meant eliminating (see p 66)

On Question 1)		All 0's and 5's.
On Question 2)		All "Yeses"
On Question 5)	a) b) c)	Those who had "yes" for all three.
On Question 6)		All "Yeses".

After this had been done, the r was calculated between the two tests, using the 39 subjects who remained. It came to + ·22.

The same procedure was then applied to the

maze B and card-sorting tests, eliminating from the maze test those who actively disliked the test, or liked it very much indeed; who were very much upset by being blindfolded; and who were nervous or over-cautious at the beginning of the test. In the card-sorting test, due to so many people liking the test very much indeed, this was not regarded as being a peculiar mode of reaction to the test, and these people were not eliminated, but those who "didn't like it very much", as well as those who"actively disliked" it were eliminated. Using the 44 subjects who remained, the correlation between the two tests came to +·13.

Correlations were then worked out between each of the various tests, using only the 23 subjects who remained in all four tests, on the grounds that these people seemed to be least affected by emotional factors. If the correlations were raised, when using only these subjects, it would suggest that learning ability might be a general factor, but that it was obscured when making large scale surveys, and using all subjects' results indiscriminately, by variations deriving from varying emotional and attitudinal factors. The correlation, coefficients calculated, using the twenty-three subjects who remained, are those shown in Table 14.

TABLE 14.

	M vs.C.	M vs.S.	M vs.N.	S vs.C.	S vs.N.	C vs.N.
Using 23 S's	·09	·33	·15	·41	·27	·35 Av·27
Using 70 S's.	·05	·25	·02	·18	·34	·13 Av·16

Except for the correlation between the substitution and nonsense syllable tests, the coefficients have been raised in every case. The greatest increase is

that found in the correlations between the substitution and card-sorting tests, which has been raised from +·18 to +·41. However the difference between the two r's does not reach the 5 per cent level of significance $\frac{x}{PE}$ being 2·07, instead of 3 (Shepherd Dawson 14 pg.138).

In spite of this, the fact that all the correlations, except one, have been raised by applying the foregoing procedure, and that the mean has been raised from ·16 to ·27, does seem suggestive, and tends to support the theoretical reasons for which the elimination technique was applied. One cannot infer from these results that a general learning ability of any importance exists, but one can argue that the extremely low correlation found in the past, may have been due in part to emotional factors exerting a differential influence in the learning situation, and that this might be worth investigating more precisely than has been done here.

H. <u>Elimination of subjects, with differences in affective attitude.</u>

It was decided to eliminate subjects specifically on their likes and dislikes of the test. It was felt that if a subject liked one test very much indeed, and disliked another, this might interfere with her relative ability on the two tests, and a number of such subjects might be lowering the correlation coefficient, due again to their emotional attitude, rather than to differences in learning ability.

Perhaps the criticism will be levelled that the subjects who disliked a test had this

affective attitude towards it, due to the fact that their learning ability on the test was low. As the questions were asked after the 1st. trial, and again after the 10th. trial, this criticism would apply only to the answers received after the 10th. trial. After the 1st. trial the subjects have little idea as to what their learning ability for that task will prove to be. The fact that they feel they dislike the test, may hinder them from learning the task to the best of their ability, but their inability to learn the task will not influence the answer they give after the first trial. For this reason it was the first set of answers that was considered.

Each subject was asked after the first trial, whether she

(a) liked the test very much indeed ... 5 points.
(b) fairly well ... ... 4 points.
(c) moderately ... ... 3 points.
(d) not very much ... ... 2 points.
(e) not at all ... ... 1 points.
(f) actively disliked it ... 0 points.

She was rated according to a 6-point scale; if she answered (a) she obtained 5 points, (b) 4 points, (c) 3 points, (d) 2 points, (e) 1 point and (f) 0.

In eliminating people according to their likes and dislikes of two tests, everyone was eliminated who differed by more than one point, on the answers given to the first asking of the questions. For example, if a subject liked one test "fairly well" (4 points), and the other test "not very much" (2 points), she was eliminated. If however she like the one test "fairly well" (4 points), and the other moderately (3 points) or "very much indeed" (5 points), she was retained. Of course all people

who gave the same answer for both tests were not eliminated.

Perhaps a qualitative analysis of some of the subjects' results, which were eliminated by this procedure will explain:

(i) How such people, due to factors other than differences in learning ability, can upset the correlations;

(ii) certain deficiencies in using time difference as the criterion for measuring learning ability, and

(iii) the unsuitability of mazes as tests of human learning ability.

Considering Maze B - people who dislike the test, seem to get either poor results, due to the fact that their time does not decrease much after the second trial, or they may get good results due to the fact that the time taken for the first couple of trials is so long that, even without getting to know the correct path through the maze much better, they are able to reduce their time considerably.

Compare the results of the following two subjects, who both said they "actively disliked" the test.

TABLE 15. RESULTS OF TWO SUBJECTS WHO ACTIVELY DISLIKED MAZE B.

A.	1.	190·0	B. 1.	105·5
	2.	30·0	2.	115·5
	3.	61·4	3.	177·6
	4.	64·2	4.	23·9

TABLE 15. (Continued)

A.	5.	236·5	B.	5.	53·6
	6.	93·3		6.	76·8
	7.	30·4		7.	59·3
	8.	42·2		8.	17·2
	9.	74·8		9.	35·4
	10.	33·9		10.	40·0

A gets a TD of 0, as the second trial happens to be her shortest, while B gets a TD of 98·3, although from observation it was noted that neither subject was very confident, or had a very good idea of the maze at the end of ten trials. This shows that TD is an unreliable criterion by which to judge learning ability.

Of the 0's, 1's and 2's (i.e. people who disliked the test) eliminated from Maze B.

A.) 4 were of the A type,(13·0; 19·0 and 18·2) getting low TD's.

B.) 6 were of the B type,(252·8; 215·7; 140·2; getting high TD's.(98·3; 101·6; and 116·9)

These A & B types of subjects have all got scores well below and above the mean (47·5) for the Maze Test. Comparing their results on this test with those obtained on the substitution test, we find that on this test all the subjects have TD's near the mean (45·4), therefore it can be seen how these subjects will tend to lower the correlation coefficient between the two tests. If their extremely high or low TD's on the Maze are due to their dislike of the test, and not due to lack of learning ability, surely we are justified in eliminating them.

C.) Two other people eliminated were noted as being poor learners, obtaining fluctuating curves, and not

really knowing the maze by the end of 10 trials. These had TD's of 48·4 and 61·5

The writer would class these three types of people as poor learners, although some had high, some average and some low TD's.

D.) 3 subjects, though from TD's alone one would not say so, she would class as good learners, although they disliked the test. These subjects all knew the maze well by the end of ten trials, and were tracing it quickly. Their TD's were 66·0, 32·5 and 20, as compared with the mean (47·5), which shows that 2 would be considered below and one above average in learning ability.

E.) The four subjects who "liked the test very much" all obtained low TD's, yet the writer would class them all as good learners. The low TD's can be accounted for by the fact that after the first trial they had such a good idea of the maze, that it was very hard for them to improve much upon the time taken to do trial 2. The TD's obtained by these subjects were 31·8; 26·5; 12·3; 10·0. Comparing these results with those obtained on the substitution test, one subject got well above the mean on this test i.e. 86, the others were all slightly below. Such individuals probably do not lessen the size of the correlation coefficient as much as the A & B types.

The remaining eleven people eliminated all got 3's and 4's. One of the eleven had a very long 2nd trial, and consequently a TD of 255·5, but the other results were not exceptional.

From this analysis of results, it would seem that like and dislike of the tests, did not go

consistently with high or low TD. Dislike of the test went more consistently with learners, whom, on a qualitative basis, were classed as poor, and like of the tests with learners qualitatively regarded as good. High and low TD's do not go consistently with good or poor learners, which shows either the unreliability of the time difference criterion of judging learning ability, or the unsuitability of the maze as a test of human learning ability.

To see which of these two alternatives is mainly responsible for the inconsistency, consider the Substitution Test B. The same type of thing occurs here, but not to such a marked degree. Only two of the subjects who disliked the test obtained high TD's, due to having exceptionally long second trials. Mostly the TD's of those people who disliked the test were small, but comparing the results with TD's of people getting affective scores of 3, 4 and 5 (i.e. who liked the test), there seems to be little difference. There is one person comparable to the E class of maze learners, who, liking the test very much, took only 87·8 seconds for the second trial. She reduced her time to 68·4 seconds, which is unusually low, but this only gave her a low TD of 19·4.

As the same kind of inconsistencies occur in the substitution test results, it would seem that the unreliability of the time difference method is partly responsible, but the greater frequency of such inconsistencies in the maze results, perhaps points to the undesirability of using a maze as a learning task.

In the card-sorting test the large

majority of people were not able to reduce their times very much, and no exceptional trends were noticed in the results of those people eliminated from this test.

In the Nonsense syllable test B most people had very similar times for their second trials, as it seems very hard to improve much on this test during the early trials, and it is also difficult to spend an exceptionally long time on any particular trial, therefore those people who disliked the test tended to have short D's, and those who liked the test had large D's as one would expect. Again a few subjects could be classed as good learners in spite of the fact that they disliked the test, and therefore they had shorter TD's.

Returning to the correlation coefficients obtained by this elimination technique, they were found to be those shown in Table 16.

TABLE 16

Elimination technique.

M vs. S.	M vs. C.	M vs. N.	S vs. C	S vs. N	C vs. N
+ .50(40 s's)	+ .24(43 s's)	-.06(47 s's)	+ .13(47 s's)	+ .34(42 s's)	-.17(29 s's)
Using 70 s's					
+ .25	+ .05	+ .02	+ .18	+ .34	+ .13

(Using this technique the same S's are not used in the correlations of the different tests as they were in the previous method.)

From these results it will be seen that using only those subjects who were similar in their likes and dislikes of the test, makes quite a large difference to two of the coefficients. It raises the r between Maze B and substitution Test B from .25 to .50. The difference between the two correlations

does not reach the 5 per cent level of significance, as $\frac{x}{PE}$ = 2·27, instead of 3 (Shepherd Dawson. 14 pg.138). It also raises the r between Maze B and card-sorting from .05 to .24, but this rise is not significant either. The other correlations all become lower, or remain unchanged. The coefficient between cards and nonsense sullables is changed from +·13 to -·17. Why is this correlation so considerably lowered? Firstly, only 29 subjects were left in for this correlation, and statistically this number is too small a one on which to base any conclusions. The second reason that can be suggested for the lowering of this correlation is that subjects who liked doing the card-sorting test improved so much between trials 1 and 2, that little improvement showed up between trials 2 to 10, and hence small TD's were obtained. On the other hand the nonsense syllable test is the test in which this type of behaviour occurs least frequently, and therefore a subject who liked the test, would probably still have a fairly high time score during the second trial, which can subsequently be considerably reduced and a large TD obtained. Many people did react to the two tests in the manner outlined above and the writer considers this partly accounts for the fact that the correlation was lowered by using only subjects who differed by 1 or less, on their likes and dislikes of the tests. Those who did not like the first test, often had high TD's on it; and often also high TD's on the second test. Eliminating these left a greater proportion of those with low TD's on one test and high TD's in the other.

From considerations such as this, it was felt that too many people were being eliminated by this rather stringent method. It was therefore decided to eliminate only those people who differed by more than 2 points. For example, if a subject liked one test "very much indeed" (5 points) and the other "not very much" (2 points), "not at all"(1 point) or "actively disliked it" (0 points), she was eliminated. If she liked the other test "moderately" (3 points), "fairly well" (4 points) or also liked it "very much indeed" (5 points), she was retained.

Working on this new basis the following correlations were found.

TABLE 17.

Elimination technique

M vs. S	M vs. C.	M vs. N.	S vs. C.	S vs. N.	C vs. N.
+.36(58 s's)	+.10(62 s's)	+.06(61 s's)	+.20(60 s's)	+.27(56 s's)	+.12(53 s's)
		70 S's			
+.25	+.05	+.02	+.18	+.34	+.13

From this it will be seen that using subjects who differed by 2 or less, on their likes and dislikes of the two tests compared, there is a more general tendency for the correlations to be slightly raised. Again the greatest difference is made in the correlations between Maze B and substitution B, and between Maze B and Card-sorting, but this difference is not as great as that found when using subjects who differed by one or less. The correlations between maze and nonsense syllables, and between substitution and card-sorting, are raised very slightly (by .04 and by .02), while those between nonsense syllables and substitution, and between nonsense syllables and cards are lowered slightly (by .04 and .01).

No very definite conclusions can be drawn from these results. Using subjects who have a similar attitude of like or dislike towards two tests does tend to raise the correlation coefficient in some cases, and only in one case has it been appreciably lowered. Differences in like or dislike of the tests may not be the main factor in lowering correlations between learning tasks; perhaps a factor such as motivation makes a greater difference to the results. Motivation is a factor which Hall (23. pg.194) believes may be of prime importance. He concludes his article by saying: "the writer believes that a general learning ability of some importance might be discovered, were we able to control the differential influence of motivation and previous practice in the learning situation."

I. Elimination of subjects with differences in Motivation.

For this reason it was decided to see what the effect on the correlations would be, Of eliminating those subjects who were differently motivated on the two tests being correlated. This was done by eliminating those people who answered "yes" for Question 4. (Were you keen on doing well on this test, or didn't you care particularly whether you did well or not?) on one test, and "no" on the other. Those people who answered "yes" on both tests, or "no" on both tests were kept in. As it may again be argued that the subject's statement at the end of the test that she was keen to do well on the test might have been due to the fact that she had been doing well, the answers to Question 4 a) only

were the ones considered - i.e. the answers referring to the subjects' motivation at the beginning of the test.

It was found that the people eliminated by this procedure, were not necessarily those people who had differed on their likes and dislikes of the test, see Table 18. In this Table column A gives the numbers of the subjects eliminated by the motivation technique. Column B gives the numbers of these subjects, who were among the people eliminated on the grounds that they differed by more than 1, on their likes and dislikes of the Test. Column B gives the numbers of subjects from column A who differed by ½ or less than 1 on likes and dislikes of the tests.

TABLE 18. TESTS BEING CORRELATED.

	Col.A.	Col.B.	Col.C.
	Subjects eliminated by motivation.	No. of A eliminated by attitudes.	No. of A retained by attitudes.
Cards vs.Sub.	20	11	9
Maze vs.Non.	23	9	14
Maze vs.Sub.	17	5	12
Maze vs.Cards	14	4	10
Sub. vs.Non.	24	8	16
Cards vs.Non.	21	11	10

Therefore it would seem that motivation at the beginning of these tests was not intimately connected with the subjects' like or dislike of the test and that one should consider "motivation" and "affective attitude" towards the test as two distinct factors in the learning situation.

The correlations found between the various tests using only S's who were similarly motivated, are shown in Table 19.

TABLE 19.

Elimination technique.

M vs. S.	M vs. C.	M vs. N.	B vs. C	S vs. N.	C vs. N
+.33(53 s's)	+.08(56 s's)	+.07(47 s's)	+.09(50 s's)	+.44(46 s's)	+.15(49 s's)
		Using 70 s's			
+.25	+.05	+.02	+.18	+.34	+.13

These correlations show a fairly consistent rise on the original correlations, although none are raised significantly. Only one r is lowered and that by +.09. The others are raised on an average by +.05. It is interesting to note that the correlation between the substitution and nonsense syllable tests has been raised by .10, as this correlation had not been much affected by the previous three methods of elimination. This would suggest that differences in different factors may have varying effects on the intercorrelations found between the results obtained in different test situations. Using a certain two tests it may be differences in motivation which are masking the correlation which exists between the learning abilities of subjects on the two tests; using two other tests, differences in like and dislike of the tests, may be the chief factor.

J. General Discussion on the results found by the elimination techniques.

Other investigators have been surprised at the low range of correlations obtained among various tests, and the fact that theorists assume that there is a general learning ability, caused the

writer to wonder if certain other factors might not be responsible for the low correlations, and to try to eliminate these factors. One might do this by trying to get all people equally motivated and to like the tests equally, etc., but another method of eliminating such factors is by eliminating the people in whom differential factors were operative during the learning situation. For this reason introspective reports in the form of answers to set questions regarding various reactions of the subject to the testing situation, were obtained, and it was purely by regarding answers to these questions that any elimination was made. An elimination technique which took into account the results obtained by the subjects, would be unjustifiable. A Summary of the results obtained by the various elimination techniques, follows in Table 20.

TABLE 20. CORRELATIONS OBTAINED FROM THE VARIOUS ELIMINATION TECHNIQUES.

Tests	70 s's	23 s's after first elimination.	S's differing by 1 or less on likes.	No.of S's left in.	S's differing by 2 or less on likes.	No.of S's left in.	S's having same motivation	No.of S's left in.
M vs.S	.25	.33	.50	(40)	.36	(58)	.33	(53)
M vs.C	.05	.09	.24	(43)	.10	(62)	.08	(56)
M vs.N	.02	.15	-.08	(47)	.06	(61)	.07	(47)
S vs.C	.18	.41	.13	(47)	.20	(60)	.09	(50)
S vs.N	.34	.27	.34	(42)	.27	(56)	.44	(46)
C vs.N	.13	.35	-.17	(29)	.12	(53)	.15	(49)
Mean	.16	.27	.16		.18		.19	

It would appear from Table 19 that taking into account as many disturbing factors as possible, and eliminating those subjects who were affected by these factors, makes the most difference to the

general trend of the correlations. When this procedure was applied, all the correlation coefficients, except one, were raised and although the new correlations obtained were not significantly different from the others, some of the correlations were raised as much as .13 and .12. The average correlation was raised from .16 to .27.

A slight but fairly consistent trend was also noticed in the correlations obtained, using subjects with similar motivation. All the correlations, except one, were raised although again not to a statistically significant degree. The average correlation was raised from .16 to .19. This would seem to show that differences in motivation do tend to lower intercorrelations between learning abilities, but this factor does not seem to be as important as Hall believed it would be.

The results obtained when eliminating people who had different attitudes of like and dislike towards the test, are rather more diffcult to interpret. Two of the correlations are quite appreciably raised, but the effect it has on the other correlations is negligible. The qualitative analysis of the type of people eliminated from the maze gives one some idea of why the two correlations that are between the maze test and something else, are the ones that are considerably raised. The same type of qualitative analysis helps one to understand the lowering of the correlation coefficient between cards and nonsense syllables.

Throughout this investigation the need for qualitative treatment of data has struck the writer

very forcibly.

The chief difficulty encountered in this investigation, using the elimination technique, is that so few subjects are left after it has been applied, that the correlations obtained cannot be regarded as statistically reliable. However a small number of subjects tends to lower a correlation coefficient, and the fact that in nearly every case the correlations found after elimination were higher than the original ones, must indicate that the factors which we were trying to eliminate, by eliminating the people in whom they were operative, are real factors which tend to mask the correlations that exist between various learning abilities.

K. Summary.

1. The mean correlation coefficient of the 28 obtained was +.06; the mean for the 6 obtained on the facilitation tasks was +.02, and the mean for the Achievement tests was +.16. It was decided to see what the effect would be on the correlation coefficients of eliminating people in whom differential emotional attitudes were operative at the time of the learning situation.

2. By eliminating all those subjects who had an extreme attitude of like or dislike towards the test; who were upset by the type of material used in the test; who were nervous or over-cautious; and who found that doing another test previously had been a hindrance, all the r's except one were raised; and the mean correlation was raised from .16 to .27.

3. Eliminating those people who differed in

their attitudes of like or dislike towards the tests being correlated, affected two of the correlations appreciably. Using subjects who differed by 1 or less on their answers to question 1 a), the mean correlation was unchanged; using subjects who differed by two or less, the mean was raised from .16 to .18.

4. Eliminating those subjects who were differently motivated on the two tests being correlated, raises all the r's except one, and raises the mean from .16 to .19.

5. Motivation and emotional attitude towards the tests does not seem intimately connected, as the people who were eliminated on the grounds of differences in motivation, do not coincide with those eliminated because of differences in affective attitude.

6. One cannot infer from these results that a general learning ability of any importance exists, but one can argue that the extremely low correlations found in the past, may have been due in part to factors such as emotional attitude and motivation exerting a differential influence on the results, since nearly all the correlation coefficients have been raised by our attempt to eliminate people in whom differential factors of this nature were operative during the learning situation.

CHAPTER VIII.

GENERAL CONCLUSIONS, RECENT WORK AND FUTURE DEVELOPMENTS.

A. General Conclusions.

As there is a general assumption made by theorists that learning ability is a general and not a specific factor, experimental evidence on this matter seems well worth obtaining. On account of the low intercorrelations that had been obtained between various learning abilities by early investigators and more recently by Hall (23) and Husband (30), the present piece of research was undertaken to find out whether a repetition of such an investigation would give analogous results. The chief difficulties in the way of getting conclusive evidence on this subject are :

(i) That an individual's performance on a given learning task, is a function of many other factors besides learning ability. Chief among these are the objective conditions, volitional attitude and motivation of the individual, fluctuating mental and organic conditions, previous practice, positive and negative transfer of training, work methods, and the number of trials given on the specific task,

(ii) That the individual's learning ability cannot be measured directly by a single performance score, but must be deduced in some way from the set of scores obtained by

repeating the learning task a number of times.

In this investigation much time was spent in trying to devise a method of measuring learning ability from the set of ten performance scores obtained, but without very satisfactory results. It was felt that the low intercorreltations between learning abilities found in the past, may have been due to the fact that methods of measuring such ability had been unsatisfactory. A formula for measuring learning ability, which on empirical grounds seemed to take account of more factors going towards making a good learner than either the measures used by Hall or Husband, was devised. The application of the formula to the data however produced intercorrelations which were not significantly different from those obtained by using the time difference between the second and the shortest trials (this latter method is somewhat similar to Hall's method of taking time difference between the first two trials and the last two trials). The formula devised, in fact, tended to make the correlations even lower than those obtained by using the time difference method, although very slightly, and not significantly. The cause of this seemed to be that the new formula was not very well suited to covering extreme cases.

As regards the first difficulty, these extraneous factors were controlled as much as possible. The objective conditions were standardized to a certain extent by the subjects always being tested in the same room at the same time of day. The effect of previous practice was minimized by choosing tasks in which none of the subjects had had specific practice before.

Positive and negative transfer of training was guarded against by spacing the tests, particularly those of a somewhat similar nature. In spite of this spacing some negative and positive transfer of training and emotional attitude did seem to occur, particularly in the second substitution test, and to a lesser extent in the second nonsense syllable test (see Chapter VII, page117).

One of the main aims of the investigation was to find out whether tasks involving facilitation learning correlated more highly with each other than with tasks offering a greater opportunity for using achievement in the learning. When it was found that the correlations between the tasks chosen in the facilitation group were very low, it was decided to try and get some data on the disturbing effects of motivation, attitude, fatigue factors, nervousness, etc., by drawing up a questionnaire to be used when the subjects came to be tested for the achievement tasks. By using this questionnaire the investigator was to eliminate people from the two tests being correlated in whom differential factors of an emotional type were playing a part. Although the individual correlation coefficients were not statistically significantly raised by this procedure, yet the correlations as a whole were fairly consistently raised, which suggests that such factors are making a difference to the correlations obtained, and therefore every attempt should be made in any future investigation of this sort to control the motivation and other emotional factors operative in the subjects being tested.

The <u>main conclusions</u> that can be drawn from

this investigation are that :

(i) The trend of evidence seems to suggest that learning abilities as measured in this investigation and in previous investigations are not highly correlated. The evidence for this arises from the fact that only two out of twenty-eight correlation coefficients found in this investigation are above +·30, and significantly different from zero. Eight out of ninety-one were greater than +·30 in Husband's investigation; twenty-nine out of the eighty-four quoted by Hall from early investigators were greater than +·30, and two out of six of his own crude r's.

(ii) It would seem that correlations are higher among learning tasks involving more complex functions than among simpler tasks. This is borne out by a consideration of the coefficients obtained between motor, mental and ideational tasks, both in this investigation and in Husband's, as well as by a comparison of the correlations obtained for the facilitation tasks, with those obtained for the achievement tasks. These differences are not statistically significant, but they seem to show a certain trend.

(iii) The lowness of the present coefficients may be in part due to factors such as emotional attitude and motivation exerting a differential influence on the results. Again the coefficients were not raised significantly by the elimination techniques which were used to

try and find out how great an effect such emotional factors exert on the results, but the fairly consistent raising of the co-efficients does seem suggestive.

Hall (23) and Husband (30) have tended to conclude that the low correlations show absence of a general learning ability. On the Spearman (53) theory, this does not necessarily follow. According to him it is the relationship between intercorrelations that determines the presence or absence of general factors, not the mere lowness of the correlations.

In order to find out whether a general factor could be discovered, Spearman's criterion of tetrad differences was applied to the intercorrelations obtained between the achievement tasks, as the largest correlations found occurred between these tests and the coefficients were all positive. The median of the tetrad differences all taken as positive comes to ·0255, while the theoretical p.e. (53 page xi Formula 16A) is ·02732. (See Appendix V). The median observed tetrad difference for the table of correlations is less than the probable error, therefore, according to Spearman, this proves the existence of a "g" factor and "s" factors in the learning abilities used on the four tests being correlated. However the amount of such a common factor would probably be very small, due to the lowness of the correlation coefficients being used, the mean of which is only +·16.

In the case of a single tetrad difference, the tetrad difference should be less than five times its P.E., calculated according to formula 16 (Spearman 53). Taking one of our tetrad differences, namely :

	Sub. B.	Cards.
Maze B.	·25	·05
Non. S.B.	·34	·13

we get the tetrad difference = ·025, and the P.E. = ·032. This satisfies Spearman's criterion for a single tetrad difference, as the tetrad difference is less than five times the P.E., so that according to Spearman there is a common factor running through the four abilities.

However Thomson (56 pg. 149 onwards) has levelled certain criticisms against Spearman in this connection. Spearman has assumed that the tetrad differences are zero in the whole population. In actual examples what Spearman's technique does, is decide whether the actual tetrad difference found is compatible with the true tetrad difference being zero. As the actual tetrad difference will very rarely be exactly zero, it is assumed that it deviates from zero, due to sampling errors. Its theoretical probable error is therefore worked out by formula 16, and if the tetrad difference is found to be less than five times the p.e., it is assumed that the actual tetrad difference is compatible with the true tetrad difference being zero. Thomson points out that it is also necessary to consider whether the sample is incompatible with the opposed hypothesis that the true tetrad difference for the whole population is not zero. In order to do this we can take a region round zero which for practical purposes we are willing to accept as zero. Thomson takes ·05 as the discrepancy from zero which he is willing to accept. He holds that it is not a very rigorous demand to make, that the tetrad

difference observed should be incompatible with a hypothesis that the true value is greater than ·05, before we definitely admit the theory that it is really zero.

This means that for a single group of four correlations the tetrad difference plus 4½ times its p.e. must be within the limit of ·05, if we are to accept the existence of a general factor running through the four abilities, along with individual specific factors. Unless it is within this limit it is compatible with the hypothesis that the true value is zero, and also compatible with the hypothesis that it is not zero, and hence the existence of one common factor is not proved; all that is proved according to Thomson is that there the chances against the Spearman hypothesis are not more than 1000 to 1 (Thomson 56 page 151).

Applying Thomson's criterion that the tetrad difference plus 4½ times its p.e. must be within the limit of ·05, to the tetrad difference quoted above, we find that

(P.E X 4½) + Tetrad difference = (·032 X 4·5) + ·025

= ·169,

and Thomson's criterion is obviously not fulfilled, and on his view the facts do not prove that one common factor exists between the learning abilities exhibited on the four tests (viz. Substitution Test B, card-sorting, Nonsense-Syllables B and Maze B).

Guilford (20) maintains that Thurstone's method of factorial analysis is the most reliable. However due to the great amount of calculation involved in this method, it was decided not to apply it to the results. Furthermore our correlations are all so near

their probable errors, and in the majority of cases do not differ significantly from zero, that although the importance of employing factorial analysis is realised, it was felt that little could be gained by applying it to the present results.

In the following section it will be seen that Husband has recently tried applying factorial analysis to his results.

B. Recent Work on the Problem.

Since the commencement of this study two articles reporting further investigations carried out by Husband (31, 32) on this subject have appeared in the Journal of Genetic Psychology. In the first article (31), Husband deals with the effect of length of tests upon intercorrelations. In order to check the possibility that shortening the tests had been the cause of the low correlations obtained in his first investigation, he chose six of the original tests and increased them to four times their original length. In spite of quadrupling the length of the tests, the median correlation is only +20, as compared with the median of +13 obtained originally. This figure is a trifle higher than the median for the original series, but so little higher that one would hesitate to suggest that quadrupling the length of the tests had produced any material differences. However the fact that the median coefficient has been raised, points to the desirability of using longer tests than Husband did in his original investigation. Perhaps quadrupling the tests (which meant giving about twenty trials instead of five)) increased the length of the tests too greatly, so that the learning curves were flattening

out too much towards the end of the practice period.

In the second article (32), Husband deals with the effect of age and spread of intelligence on the intercorrelations. It has been suggested that one of the reasons for the extremely low intercorrelations obtained between difference measures of learning may be the fact that most tests have been performed on college students who represent a homogeneous population, especially in terms of intelligence, but also in range of talent and education.

A narrow range of talent operates to reduce any correlation coefficient, because differences in sheer ability are so small that personality factors such as perseverance, enthusiasm and conscientiousness; chancing on an efficient mode of attack and slight differences in experience, can materially alter one's position within the group. If the spread of abilities happened to be wider such minor factors cannot make up for discrepancies in true ability.

To see what difference a wider range of talent would make on the intercorrelations, Husband obtained thirty children each from two of the junior high schools in Madison, one the university operated school and the other a school located closer to the business and industrial areas. The total I.Q. range was from 81 to 162, with a median of 106·5. The children were in the seventh and eighth grades (the state laws necessitating that virtually everyone will still be in school in these grades) and the age range was from 11½ to 16½ years, with a median of 15.

The tests used were six from Husband's original battery, which were deemed especially suited

for junior high school students.

The median correlation of the fifteen correlations obtained from these tests, was +.10, which is virtually insignificant.

Husband considered the possibility that a higher range of intelligence might have produced a greater differentiation among subjects of different degrees of innate ability which would mean that even if learning and aptitude were not perfectly correlated, at the same time the best learners may have been above average in all performances, and the poor learners below average. Such was not found to be the case.

Most of the highest correlations appeared between learning tasks and intelligence. The highest single coefficient was between ideational memory for a reading passage and intelligence, being +.52. Correlations involving the more motor tasks either among each other or between themselves and complex learning or intelligence, in general ranged much lower.

From the results found on these two further studies Husband concludes once more that learning abilities are specific rather than general.

Clyde Coombs and Husband have an article, now in press, which gives an account of the application of factorial analysis to Husband's original table of intercorrelations. About this Husband writes (31) : "in spite of the low range of correlations, at least four more or less general factors presented themselves. In addition there may be a number of special abilities." Unfortunately this statement is not elaborated,

therefore the method of factorial analysis applied and the nature of the general factors is not known at present.

C. Future Developments.

Husband's finding of four more or less general factors, leads the writer to stress the fact that the time is not yet ripe for forming the conclusion that learning abilities are specific rather than general. There are many aspects of the problem that are still in need of intensive study, and which, if investigated thoroughly, may show that learning ability is more general than it would seem to be from this and previous data on the subject.

An extensive investigation should be carried out on learning tasks alone, so that tasks could be found which do not give such fluctuating curves as were obtained on the maze tests, for example. The tasks should also be of such a nature that previous practice plays a minimum part in determining the subjects' scores. In order to find such tests much ingenuity will have to be used, and much patient experimentation is necessary.

Although the results obtained by trying to devise tests requiring more facilitation learning and others involving more achievement learning, do not seem to have been particularly fruitful, yet this approach to the study of learning, may be usefully developed further in future investigations. An interesting method of studying learning according to the process required in the particular learning situation, may be to take two learning tasks which

are virtually the same (such as two nonsense-syllable mazes, using different syllables in each but otherwise drawn up in exactly the same manner), and see whether the results obtained from two such tests correlate highly. If this proves to be so, then build up a different test by introducing an opportunity to educe relations and correlates, say, and see whether the test continues to correlate highly or whether, by the introduction of an opportunity to use a different process in learning the material, the correlation becomes much lower. If such an introduction did not make a great deal of difference to the correlations between two mental tasks, would it make a greater difference in motor tasks? One could perhaps take two very similar directional mazes (or perhaps the same maze learned from different ends) and see whether the results correlate highly. Then change the maze so that there are exactly the same number of passages and of the same length, but arranged in a linear fashion, so that the subject can no longer educe directional, and fewer spatial, relationships. Would the results obtained on such a maze still correlate with those found for the original maze?

The main difficulty in such an investigation would be the negative and positive transfer of training that would be likely to occur. If sufficient time elapsed between the two testing periods, this effect may be lessened. Such a procedure may also yield data useful for finding the <u>reliability of a learning task.</u> A maze learnt from both ends involves the same distance to be covered by the subject and the same

amount of spatial relation eduction, therefore if all the subjects' performances on the maze learned from different ends correlate highly, the test could perhaps be regarded as reliable and the correlation coefficient thus obtained may be used as the reliability coefficient. The learning situation might be very different in the reversed process, so that a low correlation coefficient would not prove unreliability. This vexed question of finding reliability coefficients is an important one as inter-test correlation coefficients cannot be corrected for attenuation unless the reliability coefficients of the test material are known.

Once suitable and reliable learning tasks had been devised, the next problem would be to devise a better method of measuring learning ability from the scores obtained on the test material. A few tentative suggestions on this topic have been made at the end of Chapter VI. If the test material is such that a large number of trials can be given, the common points of mastery technique may be applied to advantage. Some investigation into measuring learning ability by retentivity may prove fruitful. As this will ensure that subjects pass through a final common point of mastery, such an investigation may be combined with one studying the common points of mastery technique.

Apart from this more complicated problem of measuring learning ability, which is the best unit of measurement to use on the learning task - time, number of errors, or amount done in a certain time? On theoretical grounds the present investigator thought time would be the most suitable measure, and yet cases kept occurring where it seemed as if time

had not been a wise measure to choose, as time per trial can vary so considerably that for purposes of comparison it is hard to know how to deal with the results. Amount done per testing period may give results which are not so extreme. If extremes show up here, they will be ones of efficiency, not inefficiency; that is, a person who has done a great deal in the given period, will be one showing extreme ability on the task; while a person who takes an extremely long time is one who seems to lack such ability.

One of the things which struck the writer most forcibly in this investigation is the <u>need for qualitative analysis of data</u>. People may continue to work out intercorrelations between various tasks ad infinitum, but such investigations will give us nothing but correlation coefficients. It is time that experimenters tried to get evidence from the subjects themselves on some of the factors which lie behind and cause the size of the correlation coefficients obtained. Some attempt was made at getting qualitative information from the subjects in this investigation. A much more comprehensive questionnaire could be drawn up including questions asking the subjects whether they felt they had learnt Task A more quickly than Task B, and what they thought the reason for this had been. A direct question such as this would perhaps be very useful, followed by questions asking whether the subject had liked Task A better than task B, had been keener to do well on Task A, etc. Asking questions in this comparative way will probably be more useful than merely asking the subject whether she was keen on doing well on a test or not. A

subject may be enthusiastic on both Tasks A and B, and yet may be more highly motivated on A than B; this does not show up in the direct questioning method used in this investigation.

By this method further investigations could be carried out on the effects of motivation and attitude towards the test on intercorrelations. If a large enough group could be tested and then correlations be worked out only using those people who were actually equally motivated and liked the tests the same amount etc., perhaps higher intercorrelations between learning abilities would be found.

Even if the intercorrelations found by this procedure were not very large, one could at least be fairly sure that the size of the correlations found in this way was due to the learning abilities of the subjects and not to factors extraneous to the learning process. One would therefore be justified in applying factorial analysis to the data found by this procedure, and it seems desireble that factorial analysis should be applied to intercorrelations between learning abilities before any conclusions as to the generality or specificity of learning ability can be drawn.

In conclusion it may be stressed that in spite of the low intercorrelations found between learning abilities (as measured by the various investigators) in this and previous studies, the problem has by no means yet been fully investigated. Before any conclusions can be drawn as to the nature of learning ability or learning abilities, extensive work

must be done :

(i) on devising suitable learning tasks for an investigation of this nature;

(ii) on methods of finding the reliability of a learning task;

(iii) on methods of measuring learning ability;

(iv) on the units of measurement to be used for measuring improvability on a task.

(v) on a qualitative analysis of data.

(vi) on factorial analysis of data which has been obtained in such a way that the results are due to differences in learning abilities of the subjects, not to factors extraneous to the learning process.

APPENDIX I. RANGES OF SECOND TIMES, SHORTEST TIMES AND TIME DIFFERENCES.

TABLE A. Table showing range of times taken on the second trial, range of shortest trials and time differences.

Name of Test.	Range of times for second trial.		Range of times for shortest trial.		Range of Time Differences.	
	Highest	Lowest	Highest	Lowest	Highest	Lowest
Maze A.	302.3	28-5	71-4	30-5	341-0	1-1
Block-packing	194-4	46-5	48-7	22-2	196-7	16-6
Substitution A	236-7	105-4	136-0	60-0	134-9	6-9
Non.Syllables A	221-8	70-2	117-2	35-6	136-6	11-3
Maze B	272-0	21-4	54-5	8-5	265-5	0-0
Card-sorting	76-0	26-2	34-7	17-3	51-0	5-4
Substitution B	209-3	75-0	129-3	50-6	109-2	11-7
Non.Syllables B	109-7	43-6	53-9	13-2	84-3	0-3

APPENDIX II. TABLES OF SCORES FROM WHICH THE CORRELATION COEFFICIENTS WERE CALCULATED.

TABLE B. SCORES OBTAINED ON THE FACILITATION TASKS.

No. of subject.	Using TD				Using $\frac{TD^2}{A \times I}$ X 1000			
	Mase A.	Sub. A.	Blocks	Non. S.A.	Mase A.	Sub. A.	Blocks	Non. S.A.
1	17	65	54	41	71	264	200	111
2	62	34	98	66	206	43	294	143
3	29	26	40	22	178	48	126	45
4	47	30	123	11	221	57	340	33
5	21	27	62	30	120	52	214	100
6	53	77	81	21	181	122	162	52
7	15	75	92	35	104	106	189	121
8	34	41	61	43	131	81	156	164
9	93	53	76	50	307	71	305	120
10	20	34	117	41	109	62	261	83
11	6	38	101	16	10	59	229	70
12	22	16	73	27	84	26	162	54
13	18	48	30	50	134	150	259	171
14	3	62	43	94	1	103	134	196
15	23	43	50	50	31	86	156	129
16	29	24	80	30	28	37	311	112
17	6	51	54	57	9	134	157	146
18	31	35	104	43	91	47	322	123
19	174	64	78	44	408	105	194	93
20	31	26	23	42	122	54	50	103
21	33	47	69	70	119	92	276	200
22	34	49	52	65	60	94	156	307
23	30	39	100	137	83	125	331	306
24	35	29	17	58	183	51	106	183
25	37	57	55	102	109	81	170	106
26	18	43	143	31	53	101	337	56
27	30	7	107	61	140	7	230	300
28	12	32	36	15	33	49	151	30
29	23	62	63	55	126	177	242	127
30	21	42	73	41	90	84	306	95
31	46	53	73	60	188	133	309	146
32	92	23	97	50	266	36	180	108
33	55	46	134	34	302	85	201	102
34	28	33	107	20	184	46	326	75
35	71	40	72	99	215	57	166	87
36	36	10	57	75	122	20	164	101

No. of Subject.	Using D				Using $\frac{D^2}{A \times Y} \times 1000$			
	Hand A.	Sub. A.	Blocks	Hon. S.A.	Hand A.	Sub. A.	Blocks	Hon. S.A.
37	6	7	33	121	46	9	180	218
38	50	46	108	123	300	89	396	303
39	45	13	65	73	87	13	170	135
40	5	26	38	92	31	66	178	247
41	7	40	89	34	8	46	168	133
42	7	46	23	81	30	73	44	213
43	11	31	39	89	37	46	117	98
44	17	47	67	131	80	89	174	196
45	1	74	89	24	0	119	148	74
46	20	38	54	37	30	97	173	86
47	24	35	61	89	88	71	218	146
48	40	68	83	45	148	117	98	116
49	8	89	49	70	16	99	238	209
50	36	39	26	89	100	13	99	300
51	32	135	80	68	83	171	174	116
52	30	68	89	47	139	119	235	171
53	28	62	50	41	186	150	166	101
54	9	51	65	74	13	99	193	331
55	35	60	82	91	93	109	300	192
56	6	20	35	105	53	56	128	100
57	18	42	85	92	43	73	140	151
58	27	70	72	42	66	113	366	93
59	13	36	19	89	60	36	86	117
60	241	33	61	86	403	65	233	146
61	51	36	104	62	175	32	500	104
62	61	13	67	42	306	13	230	77
63	8	54	62	49	17	109	104	102
64	8	41	71	82	21	71	302	216
65	31	25	72	43	43	46	232	82
66	9	66	45	96	31	101	173	186
67	100	60	62	75	361	170	106	180
68	27	64	145	83	89	84	321	115
69	26	77	92	64	135	123	246	170
70	16	19	35	55	119	34	250	161

TABLE C. SCORES OBTAINED ON THE ACHIEVEMENT TASKS.

No. of Subject.	TD				$\frac{TD^2}{A \times I} \times 1000$			
	Mase B.	Sub. B.	Cards.	Non. B.	Mase B.	Sub. B.	Cards.	Non. B.
1	61	53	18	39	354	136	114	156
2	50	43	31	71	207	99	196	190
3	20	35	24	62	99	97	233	301
4	35	24	20	8	103	30	137	14
5	9	19	12	37	84	43	133	90
6	17	66	16	42	101	173	100	96
7	30	65	24	30	183	154	147	73
8	10	56	22	57	17	103	123	172
9	9	30	30	55	142	67	106	133
10	98	49	14	27	189	142	142	55
11	101	47	23	40	530	110	130	131
12	117	57	21	40	634	98	156	215
13	19	16	11	50	210	70	104	161
14	256	102	26	70	372	149	190	116
15	0	19	16	31	0	58	145	224
16	25	31	11	33	306	48	100	79
17	140	53	13	49	509	117	75	127
18	25	29	19	47	309	34	71	150
19	18	46	15	69	25	107	98	281
20	10	26	20	36	39	65	142	152
21	74	57	14	75	384	87	77	224
22	27	31	20	62	215	70	174	249
23	48	28	15	49	100	102	152	148
24	31	32	14	54	153	64	196	133
25	88	89	19	72	303	303	70	175
26	20	12	8	56	14	22	46	163
27	13	19	13	65	21	44	92	301
28	216	29	10	51	603	34	72	116
29	5	30	8	82	34	111	46	376
30	66	41	12	56	383	70	95	133
31	65	45	13	28	305	111	196	72
32	253	32	14	35	471	70	163	104
33	25	38	16	38	13	155	109	73
34	62	33	13	40	118	80	88	127

No. of Subject.	TD				$\frac{TD^2}{A \times Y} \times 1000$			
	Nose B.	Sub. B.	Cords.	Mon. B.	Nose B.	Sub. B.	Cords.	Mon. B.
35	63	109	13	70	189	134	86	186
36	57	69	51	84	235	91	401	246
37	19	15	13	30	85	47	137	77
38	13	36	6	47	63	73	37	306
39	21	46	24	56	30	70	135	126
40	10	41	9	60	140	106	76	297
41	22	30	16	30	37	39	83	81
42	20	37	11	38	84	83	81	90
43	72	65	20	66	301	100	108	114
44	20	56	16	70	47	142	105	167
45	33	30	16	33	230	36	303	93
46	21	54	16	39	76	84	115	104
47	12	34	20	64	61	66	93	240
48	10	33	13	39	81	83	112	94
49	18	53	32	38	123	96	177	140
50	38	62	6	32	145	109	51	101
51	120	42	36	51	304	94	118	240
52	20	65	19	42	36	104	118	130
53	20	64	18	33	202	120	137	261
54	13	46	26	63	11	103	196	164
55	44	32	10	55	137	72	87	102
56	30	32	13	66	253	107	104	147
57	30	60	27	76	66	101	115	400
58	12	41	26	54	130	46	116	175
59	47	65	12	49	300	117	112	127
60	30	41	5	43	98	137	37	211
61	15	30	24	73	70	116	106	173
62	33	86	9	77	115	181	86	242
63	94	40	33	39	244	80	218	126
64	21	52	18	65	61	68	184	331
65	106	41	15	56	252	51	60	261
66	56	46	21	60	424	56	134	100
67	82	60	6	48	95	155	25	106
68	17	63	25	35	66	117	100	113
69	74	22	30	40	345	64	107	107
70	10	20	9	53	36	52	59	100

Sub.D. using Mo.

TABLE D. ILLUSTRATING THE CONSTRUCTION OF THE PRODUCT-MOMENT CORRELATION COEFFICIENT FROM THE DATA IN A FREQUENCY TABLE.

Mom.S.(D) using Mo.

	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	δy	Δy	$\delta y\Delta y$	$\delta y\Delta y^2$
100-110								11 12		2	6	12	72
90-100										-	5		
80-90				1 4				11 8		3	4	12	48
70-80										-	3		
60-70				111 6	1111 8		1 2	1 2	1 2	10	2	20	40
50-60				111 3	11 2	1 1	1 1	111 3		10	1	10	10
40-50			11 0	1 0	11 0	1 0	111 0	1 0		15	0	0	
30-40				111 -5	11 -2	111 -3	1111 -4		1 -3	15	-1	-15	15
20-30	1 -2		1 -2	11 -4	111 -6	11 -4				9	-2	-18	36
10-20				111 -9		111 -9				6	-3	-18	54
δx	1	-	3	16	13	15	9	9	2	70		3	275
Δx	-4	-3	-2	-1	0	1	2	3	4			$\Sigma fxdy$	$\Sigma fydy^2$
$fxdx$	-4	-	-6	-16	0	15	18	27	8	40	$\Sigma fxdx$		
$fxdx^2$	16	-	12	16	-	15	36	81	32	210	$\Sigma fxdx^2$		
$\Sigma' fxydy$	-2	-	-2	-5	2	-15	-1	25	1	3	Σfxy		
$\Sigma(\Sigma' fxydy)$	8	-	4	5	0	-15	-2	75	4	79	$\Sigma fxy\Delta x\Delta y$		

For further calculations see page 160 .

APPENDIX III.

Method of Calculating Correlation Coefficients

Found by Method outlined by Holzinger "Statistical Methods for Students in Education" (Ginn & Company) page 151.

For calculating a correlation coefficient for a distribution table, the following formula is used :

$$r = \frac{\Sigma f_{xy} d_x d_y - \frac{(\Sigma f_x d_x)(\Sigma f_y d_y)}{N}}{\sqrt{\left[\Sigma f_x d_x^2 - \frac{(\Sigma f_x d_x)^2}{N}\right]\left[\Sigma f_y d_y^2 - \frac{(\Sigma f_y d_y)^2}{N}\right]}} = \frac{a}{\sqrt{bc}}$$

The complete calculation with this formula is illustrated in Table D. Appendix II for the data obtained from Nonsense Syllables (B) and Substitution (B)

$$a = \Sigma f_{xy} d_x d_y - \frac{(\Sigma f_x d_x)(\Sigma f_y d_y)}{N} = 79 - \frac{(40)(3)}{70} = 77.3$$

$$b = \Sigma f_x d_x^2 - \frac{(\Sigma f_x d_x)^2}{N} = 210 - \frac{(40)(40)}{70} = 187.1$$

$$c = \Sigma f_y d_y^2 - \frac{(\Sigma f_y d_y)^2}{N} = 275 - \frac{(3)(3)}{70} = 274.9$$

By four-place logarithms,

log b	= 2·2720	log a	= 1·8882.
log c	= 2·4392	log $\sqrt{\text{prod.}}$	= 2·3556
log prod.	= $\overline{2}$|4·7112	log r	= $\overline{1}$·5326
log $\sqrt{\text{prod.}}$	= 2.3556	$\therefore$ r	= +·3409

The computation down to the values $\Sigma f_x d_x^2$ and $\Sigma f_y d_y^2$ is the same as for the standard deviation, so that the values for b and c may be readily obtained.

The calculation for a presents a little more difficulty. The quantity $\Sigma f_{xy} d_x d_y$ is the result of multiplying each cell frequency by its dx and dy and then adding all the products so formed.

A more convenient method of calculation, however, is to multiply the cell frequencies in a particular column by the appropriate dy values, add the results found, and multiply this sum by the d_x value of the column. Continuing in this way for all the columns and adding the products thus found gives the required $\Sigma fxy\, d_x d_y$.

The symbol Σ has been used to indicate summation over the whole table, i.e. over N items. In order to distinguish summation over the columns, this has been designated in the table by Σ'. Thus, $\Sigma' fxy\, dy$ means the sum for one column of fxy multiplied by the corresponding values dy.

A useful check on the computation of a is shown by the double arrow in Table D. The sum of the quantities $\Sigma' fx_y d_y$ should be the same as $\Sigma f_y d_y$, or $\Sigma(\Sigma' fxy d_y) = \Sigma f_y d_y$.

The correction $\frac{(\Sigma fxix)(\Sigma fy d_y)}{N}$ applied to $\Sigma fx_y d_x d_y$ will sometimes be positive and sometimes negative, and it should be remembered that it is to be subtracted algebraically.

When appreciably fewer subjects than fifty were being used in calculating the correlation coefficients, Pearson's product-moment coefficient of correlation was used, as outlined on page 10 in "How to Calculate Correlations" (G.H. Thomson. 55).

APPENDIX IV. **METHOD OF CALCULATING THE SIGNIFICANCE OF A DIFFERENCE BETWEEN TWO CORRELATION COEFFICIENTS.** (Shepherd Dawson "An introduction to the computation of Statistics" pg. 138.)

To estimate the significance of a difference between two correlation coefficients, find the probable error of this difference and find what multiple the difference is of its probable error. e.g. See whether difference is significant between correlations found between maze and substitution test (·25) using 70 subjects and between maze and substitution using 40 subjects, who differed by one or less on their likes and dislikes of the test (·50).

r_1 = ·25 with a P.E of ·08.

r_2 = ·50 " " P.E of ·08.

$$\text{P.E diff.of } r_1 \text{ and } r_2 = \sqrt{P.E_{r_1}{}^2 + P.E_{r_2}{}^2}$$

$$= \sqrt{·08^2 + ·08^2}$$

$$= \sqrt{·0128}$$

$$= ·1132$$

$$\frac{r_1 - r_2}{P.E_{r_1 - r_2}} = \frac{·50 - ·25}{·113} = \frac{·25}{·11} = 2·27$$

To reach the 5 per cent level of significance $\frac{x}{PE}$ must be 3, hence these coefficients are not significantly different.

APPENDIX Va) EXAMPLE OF APPLICATION OF SPEARMAN'S 2-FACTOR THEORY, TO ACHIEVEMENT TASKS.

(The correction for attenuation is not necessary - if Spearman's criterion is passed by the correlations when they are corrected for attenuation, then it must also be passed when they are not so. (Spearman 53. pg. vi Appendix).

To see whether there is a "g" and "s" factors in a battery of tests we calculate all the tetrad differences for the correlations we have.

	Maze B.	Sub.B.	Cards.	Non.S.B.
Maze B.	-	·25	·05	·02
Sub.B.	·25	-	·18	·34
Cards	·05	·18	-	·13
Non.S.B.	·02	·34	·13	-

Four tests give 6 intercorrelations, and 6 intercorrelations give 3 tetrad difference.

·05 ·18 ·25 ·18 ·34 ·25

·02 ·34 ·02 ·13 ·13 ·05

Tetrad diff.·0134 Tetrad diff.·0289 Tetrad diff.·0255

Taking all the tetrad differences as positive, we compare their median with the theoretical probable error, (calculated from formula 16A Spearman 53. pg.xi Appendix)

$$P.E. = \frac{1 \cdot 349}{N^{\frac{1}{2}}} \left[r^2(1 - r)^2 + (1 - R)s^2 \right]^{\frac{1}{2}}$$

Where $R = 3r \frac{n - 4}{n - 2} - 2r^2 \frac{n - 6}{n - 2}$

Where r = mean of all the correlations concerned.

N = number of subjects tested.

n = number of different tests used to obtain the correlations.

s = Standard deviation of correlation coefficients from r.

In this example

$$R = (3)(\cdot16)\frac{(4-4)}{(4-2)} - (2)(\cdot0256)\frac{(4-6)}{(4-2)}$$

$$= 0 - \cdot0512\left(\frac{-2}{+2}\right)$$

$$= \underline{+\cdot0512}$$

$$\text{P.e.} = \frac{1\cdot349}{\sqrt{70}}\left[\cdot0256(1-\cdot16)^2 + (1-\cdot0512)(\cdot0122)\right]^{\frac{1}{2}}$$

$$= \frac{1\cdot349}{\sqrt{70}}\left[(\cdot0256)(\cdot7056) + (\cdot9488)(\cdot0122)\right]^{\frac{1}{2}}$$

$$= \frac{1\cdot349}{\sqrt{70}}\left[\cdot0181 + \cdot0106\right]^{\frac{1}{2}}$$

$$= \frac{1\cdot349}{\sqrt{70}}(\cdot0287)^{\frac{1}{2}}$$

$$= \underline{\cdot02732} \longrightarrow$$

The median of the tetrad differences is .0255, which is therefore less than the theoretical P.E, and therefore satisfies Spearman's criterion.

b) EXAMPLE OF APPLICATION OF SPEARMAN'S TWO-FACTOR THEORY, TO FOUR CORRELATION COEFFICIENTS.

Taking the following 4 correlation coefficients

	Sub.B.(1)	Cards (2)
Maze B.(3)	·25	·05
Non.S.B.(4)	·34	·13

The tetrad difference is ·0255.

The probable error of this difference is worked out by the formula 16(53. pg.xi. Appendix), which is

$$\text{p.e.} = \frac{1\cdot349}{N^{\frac{1}{2}}}\left[r^2(1 - r_{12} - r_{34} + r^2) + (1 - 2r^2)s^2\right]^{\frac{1}{2}}$$

When r = mean of all correlations concerned.

N = number of subjects tested.

s = standard deviation of correlation coefficients from r.

r_{12} = correlation coefficient between tests 1 & 2 found from another part of the table.

In this example

$$\text{p.e.} = \frac{1\cdot349}{\sqrt{70}}\left[\cdot0361(1 - \cdot18 - \cdot02 + \cdot0361) + (1 - \cdot0722)(\cdot0123)\right]^{\frac{1}{2}}$$

$$= \frac{1\cdot349}{\sqrt{70}}\left[\cdot0302 + \cdot0114\right]^{\frac{1}{2}}$$

$$= \frac{1\cdot349}{\sqrt{70}}\left(0416\right)^{\frac{1}{2}}$$

$$= \cdot03289$$

Therefore tetrad difference of ·0255 is less than five times its P.E., and Spearman's criterion holds.

REFERENCES.

(References which are marked with an asterisk, are ones which were not obtainable, but which have a bearing on the subject, and have been referred to by other authors.)

1. * Anastasi, A. — Arch.of Psychol.1930.No.120. "A group factor in immediate memory".

2. Bartlett, F.C. — "Remembering: A Study in Experimental and Social Psychology". (Cambridge Univ.Press.)

3. Britt, S.H. — Psych.Rev.1937. 44. p.462-469 "The Learning-Remembering Process - (A reply to Prof.Cason)"

4. * Brooks, F.D. — Jr.Exper.Psychol.1924. 7. p.462-466. "Learning in the case of three dissimilar functions".

5. Carr, H.A. — Psych.Rev.1937. 44. p.274-296 "The Search for Certainty".

6. Carr, H.A. — Psych.Rev.1938. 45. p.435. "The Reliability vs.the validity of Test Scores".

7. Carr,H.A. and Kingsbury, F.A. — Psych.Rev.1938. 45. p.354-376 "The Concept of Ability".

8. Cason, H. — Psych.Rev.1937. 44. p.54-61 "The concepts of learning and memory".

9. Cason, H. — Psych.Rev.1938. 45. p.101-105 "Dr.Kellogg on the definition of learning."

10. Cason, H. — Psych.Rev. 1938.45. p.441-443 "Dr.Britt on the concepts of learning, retention and reproduction.

11. * Chapman, J.C. — "Individual differences in ability and improvement and their correlations" New York Teachers College, 1914.

12. Chappell, M.N. Psych.Rev.1940. 47. p.90-94. "The inadequacy of the Kellogg-Britt Definition of Learning".

13. Dashiell, J.F. Psych.Rev.1940. 47. p.289-306. "A Neglected Fourth Dimension to Psychological Research".

14. Dawson, S. "An introduction to the computation of Statistics". (University of London Press 1933) p.129-139.

15. Duncker, K and Krechevsky, I. Psych.Rev.1930. 46. p.176-185. "On Solution-Achievement".

16. van Dusen, A.C. Jr.of Exp.Psych.1941. 29. p.225-236. "Work Methods and Learning".

17. * Garrett, H.E. Jr.Educ.Psychol.1928. 19. p.601-613. "The relation of tests of memory and learning to each other and to general intelligence in a highly selected adult group".

18. * Gates, G.S. Arch.Psychol.1922. 58. p.74. "Individual differences as affected by practice".

19. Glaze, J.A. Jr.Genet.Psychol.1928. 35. p.225-269. "The Associative Value of Nonsense Syllables".

20. Guilford, J.P. "Psychometric Methods" (McGraw-Hill Book Company.)

21. Guilford, J.P. Psych.Rev.1940. 47.p.367-393. "Human Abilities."

22. * Gunlach, R. Jr.Educ.Psychol.1926. 17. p.387-401."The Effects of Practice on the intercorrelations of three mental tests".

23. Hall, C.S. Psych.Rev.1936. 43. p.179-196. "Intercorrelations of Measures of Human Learning".

24. * Haught, B.F. Psychol.Monog.1921. 30.No.139, 71. "The intercorrelation of some higher learning processes."

25. Heron, W.T. Comp.Psychol.Monog.1924.2. p.60. "Individual differences in ability versus chance in the learning of the stylus maze".

26. Hertzman, M. Psych.Rev.1939./p.163-175. "The Specificity of Correlations between Initial and Final Abilities in Learning".

27. Hollingworth, H.L. Arch.Psychol.1912.No.22. "The influence of caffein on mental and motor efficiency", cited by Rush (50)

28. Holzinger, K.J. "Statistical Methods for Students in Education". (Ginn and Company).

29. Hunter, W.S. "Experimental Studies of Learning" in "Handbook of General Experimental Psychology". p.554.

30. Husband, R.W. 1939 Jr.Genet. Psych./55. p.353-364 "Intercorrelation Among Learning Abilities:I"

31. Husband, R.W. Jr.Genet.Psych.1941. 58.p.427-430 "Intercorrelations Among Learning Abilities:III. The Effect of Length of Tests upon Intercorrelations".

32. Husband, R.W. Jr.Genet.Psych.1941. 58. p.431-434 "Intercorrelations Among Learning Abilities: IV. Effects of Age and Spread of Intelligence upon Relationships".

33. Jones, W. "Theory and Practice of Psychology" (Macmillan.1934)

34. Kellogg, W.N. Psych.Rev.1939. 45.p.165-184. "An eclectic view of some theories of learning".

35. Kellogg, W.N. & Britt, S.H. Psych.Rev.1939. 46.p.186-197. "Structure or Function in the definition of Learning?"

36. Kellogg, W.N. Psych.Rev.1939. 46.p.489-491. "On the Nature of Skills - A Reply to Mr.Lynch".

37. Kellogg, W.N. Psych.Rev. 1940.47. p.95 - 97. "The Superfluity of the Chappell Critique - A Reply".

38. Koffka, K. "Growth of Mind". Chap.IV and V.(Kegan Paul.1924).

39. Krechevsky, I. Psych.Rev.1932. 39. p.516 - 532. "Hypotheses in Rats".

40. Leeper, R. Jr.Genet.Psych.1935. 46.p.3-40. "The role of motivation in learning; a study in the phenomenon of differential motivational control of the utilization of habits".

41. Lynch, J.M. Psych.Rev.1939. 46.p.485-488. "A note on Kellogg's Treatment of Skills".

42. McDougall, W. "Energies of Men". (Methuen) Chap. XXIII.

43. Maier, N.R.F. Psych.Rev.1939. 46. p.241-251. "The Specific Processes Constituting the Learning Function".

44. Maier, N.R.F. Psych.Rev.1940. 47. p.43-58. "The Behaviour Mechanisms Concerned with Problem Solving".

45. Muenzinger,K.F. Psych.Rev.1931. 38.p.347-358. "The Primary Factors in Learning".

46. * Pyle, W.H. Jr.Educ.Psychol.1919. 10. p.121-128. "Is individual learning capacity constant for different types of material?"

47. Pyle, W.H. "A Laboratory Manual in the Psychology of Learning". (Warwick & York.Inc.1923)

48. * Race, H.V. "Improvability; its intercorrelations and its relations to initial ability." New York: Teachers College, 1922.

49. Ross,C.C. Psych.Rev.1936. 43. p.197-205. "A Needed Emphasis in Psychological Research".

50. Ruch, F.L. Psych.Rev.1936. 43. p.229-234. "The Method of Common Points of Mastery as a Technique in Human Learning Experimentation".

51. Seashore, R.H. Psych.Rev.1939. 46. p.123-140. "Work Methods: An often neglected factor underlying individual differences".

52. Spearman, C. "The Nature of Intelligence, and the Principles of Cognition". (Macmillan) Chap.IX.

53. Spearman, C. "The Abilities of Man: Their Nature and Measurement". (Macmillan)

54. Thomas, F.C. "Ability and Knowledge". (Macmillan)

55. Thomson, G.H. "How to calculate correlations" (George G.Harrap & Co.Ltd.)

56. Thomson, G.H. "The Factorial Analysis of Human Ability" (University of London Press.)

57. Thorndike, E.L. Amer.Jr.Psychol.1916. 27. p.550-565. "Notes on practice, improvability, and the curve of work".

58. Thorndike, E.L. "Adult Learning".Pub.May 1928 (Macmillan)

59. Thorndike, E.L. "Human Learning" (The Century Co.New York: London 1931)

60. Whipple, G.M. "Manual of Mental and Physical Tests. (Warwick and York) p.305-328 (Cancellation) p.499-515 (Substitution)

61. Woodworth,R.S. "Experimental Psychology" (Methuen) pg.160.
