CONTRIBUTION OF INTELLIGENCE TO THE LEARNING OF VOCABULARY, FACTS AND REASONING IN PSYCHOLOGY

Thesis for the Degree of M. A.

MICHIGAN STATE COLLEGE

Victor D. Sanua

1952

This is to certify that the

thesis entitled

Contribution of Intelligence to the Learning of Vocabulary, Facts and Reasoning in Psychology

presented by

Victor D. Sanua

has been accepted towards fulfillment
of the requirements for

M.A. degree in Psychology

Donald W. Johnson
Major professor

Date 11th. February 1953

CONTRIBUTION OF INTELLIGENCE TO THE LEARNING OF VOCABULARY, FACTS AND REASONING IN PSYCHOLOGY

By

Victor D. Sanua

A THESIS

Submitted to the School of Graduate Studies at Michigan State College of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Department of Psychology

1952

ACKNOWLEDGMENT

Grateful acknowledgment is made to
Doctor Donald M. Johnson for his
advice and assistance in the
course of this
research

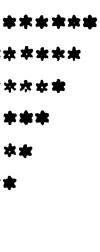

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The statement has frequently been made that intelligence is the ability to learn, but recently this statement has been questioned by several investigators. The study reported here examines the evidence obtained by correlating scores on a standard intelligence test with different measures of improvement in a college course in psychology.

The argument that intelligence is the ability to learn is based on correlations between scores on intelligence tests and scores on achievement tests in many branches of knowledge. But obviously achievement at any time depends not only on ability to learn but also on many other factors, such as previous achievement, motivation, persistence, opportunity to learn, etc. It is possible to control some of these factors, particularly opportunity to learn and previous achievement, by the use of improvement or gain scores. The subjects are tested before and after a learning period and the initial score is subtracted from the final score to get the gain score. Gain scores are then correlated with scores on an intelligence test.

Gain scores have been obtained from two kinds of experiments. Some investigators have had their subjects, after taking the initial test, practice for a standard period of time in the laboratory, then take the final test. This type of research has used such tasks as addition, subtraction and cancellation. Others have studied improvement in regular college courses.

A typical experiment of this type, which was similar to the one undertaken in this study, is by Carlson, Fisher and Young (12). Examinations based upon established facts and principles in psychology were prepared and revised by the experimenters and given as a proficiency examination. This examination was composed of multiple-choice and true-false statements. The same test was given at the beginning of the course and at the end. The experimenters made use of the scores of 118 students who had taken the Otis Intelligence test. They found that the correlation of intelligence with the pre-test was .556 while that with the post-test was .429. This is what they write in connection with this finding:

> While the scores on both pre-test and post-test correlate positively with scores on the Otis, the magnitude of the correlation is greater in the case of the pre-test than in the case of the post-test. It would appear that intelligence is more involved as a determiner of scores in the pre-test than it is in the post-test. (12, p. 30)

The next step was to correlate intelligence with the gain scores obtained by subtracting the pre-test score from the post-test score for every student. The correlation they obtained was -.039 which definitely indicates a lack of significant relation between gain scores and the Otis scores. They draw the following conclusion from their results:

> There are several possible interpretations of the results. One of these is that the proficiency test which we used was largely a memory test, and therefore did not measure increased understanding of principles. A further interpretation is that improvement following tuition is largely dependent upon non-intellectual factors. Among the more important of these, we would expect to find interest, motivation, opportunity to study, study habit, personality traits and emotional conflict at

> the time of study and at the time of taking the examinations. These factors call for attention to the importance of non-intellectual processes in accounting for differential improvement in achievement. (12, p. 33)

In spite of evidence of this kind, numerous authorities identify intelligence with the ability to learn. But with all these assertions, as Woodrow says, "no one, so far as the writer is aware has marshalled the evidence in support of the view expressed in these quotations." (41, p. 149) On the other hand numerous studies have appeared intermittently during the last three decades in the psychological journals bringing out the fact that gain scores usually do not correlate with intelligence. The results of this kind of research up to 1946 were summarized by Woodrow (41). The most pertinent results are displayed in Table I.

However, among the similar studies undertaken, it has been found that in some the gain scores correlated with intelligence. Although these correlations are rather low, they are still significant. This would naturally throw a little doubt on rejection of the notion that intelligence is the ability to learn. We shall now review some of the studies which indicate that the relation between intelligence and gain scores is still an open question.

To the best knowledge of the writer, the latest work undertaken in this field was by Simrall (27), who tackled this problem in 1947. From the popular notion that intelligence is the ability to learn, she drew a list of the implications that could be deduced from it.

TABLE I

SUMMARY OF RESEARCH CORRELATING GAIN SCORES WITH INTELLIGENCE

Experiments and Year	Achievement Test	Intelligence Test	r Initial	r Final	r Gain	Sample
Johnson 1923	Inverted reading with mirror	Army Alpha Thurstone etc.		.34	.46	60 college students
Spence 1929	Educational Psychology	Otis	.44 .34	.50 .43	.25 .09	116 college students
De Weerdt 1927	Substitution Addition Reading Cancellation Multiplication Same Opposite Multiplication by substitution	Illinois	.01 .19 .68 -.12 .06 .49 -.07	.14 .27 .63 -.14 .06 .45 -.07	.19 .28 .56 -.07 .02 .18 .20	49 fifth grade pupils
Drake 1940	Biology	A.C.E.	.34	.08	-.14	42 high school students

TABLE I CONTINUED

Experiments and Year	Achievement Test	Intelligence Test		r Initial	r Final	r Gain	Sample
McCullough 1939	Iowa Silent Reading	A.C.E.	L.	.56	.57	-.02	49 college students
			Q.	.27	.27	-.06	
	Purdue Speed		L.	.29	.27	-.07	
			Q.	.19	.23	-.08	
	Iowa Silent Reading	Cali-fornia	L.	.45	.49	.09	
			Q.	.14	.24	.29	
			L.	-.01	.18	.23	
			Q.	.15	.10	.01	
Husband 1941	Mirror Drawing					.22	100 college students
	Code Substitution					.19	
	Visual Prose Memory					.29	
	Cancellation					-.13	
	Inc. memory for code					-.15	
	Hindu-English					-.16	
Dysinger 1941	Psychology	Alpha				-.064	120 college students
Carlson	Psychology	Otis		.556	.429	-.039	118 college students

Table I continued

Experiments and Year	Achievement Test	Intelligence Test	r Initial	r Final	r Gain	Sample
				5th.	6th.	Grade 102
Woodrow 1945	Reading	Otis		.422	.169	
	Vocabulary			.431	.124	
	Arith. Problems			.508	.076	
	Arith. Fund.			.432	.012	
	English			-.065	-.028	
	Spelling			.395	.250	
Simrall 1947	Spatial Test	Otis	.604	.61	.28	95 high school students
	Perceptual Test		.59	.48	-.08	
Sanua 1952	Psychology Vocabulary	L. A.C.E.	.16	.10	-.026	100 college students
	Facts		.19	.08	-.063	
	Reasoning		.22	.27	0.080	
	Total				-.024	

She was able to destroy all these deductions with the data she obtained. However one of the shortcomings of this study is that no endeavour is made to explain the differences obtained in the correlations with the gain scores. For the perceptual test, this correlation was found to be -.079, while with the spatial test it was found to be .277. The latter coefficient of correlation would be significant at the 1% level; nevertheless no comments are made. The difference between the two correlations is almost .35 which is not necessarily negligible. Since Simrall states that there is no relation between intelligence and gain scores, it should be expected that all correlations found should be insignificant. The question which the author should have asked herself is the reason for such disparity between the perceptual and spatial test.

Connected with this difference found in the correlations of the gain scores, further data may be given that might be significant. Correlation of the initial and final scores with intelligence were found to be as follows:

	r with initial	r with final	Gain
Spatial	.60	.61	.28
Perceptual	.59	.49	-.08

In one case the correlation with intelligence remains the same, while in the other case the correlation drops. Would the difference in the gain scores correlations have any connection with this finding? All these questions will be given some interpretation at the end of this paper.

Another important point that could be raised in this experiment is whether the gain scores were the best the students could do. Simrall assumes that this is the case. This moot point is disposed of with the following lines:

> The instructions given to the subject were designed to stimulate satisfactory motivating conditions. The results of the investigation indicate clearly that the behavior of the subjects was relatively constant throughout the experiment (27, p. 32).

Such a cursory treatment of an important factor is questionable. No doubt the correlation between gain scores and intelligence is pretty low, but exceptions can be found.

The first study in Table I, by Johnson (21), differs from all the other studies since he found that the improvement scores for mirror reading correlate higher with intelligence than with the final scores, .46 and .34 respectively. From his results, he draws the following conclusion which is opposite to the one reached by Woodrow, Simrall and others:

> They show that there exists a fairly positive relation between the ability to become efficient at learning to read inverted print and intelligence as it is measured by the usual group test. It is interesting to note that it is not the absolute amount which a person reads that is most important in this connection but rather it is the rapidity of learning, acquiring new connections that is most closely related to mental ability (21, p. 541).

He furthermore illustrates his point by comparing two curves constructed on the basis of the scores obtained by 30 students above the average in intelligence and 30 students below the average. The slope of the superior group was found to be more accentuated thus

showing that the superior group gained more though the initial difference was small.

Another study, by De Weerdt (13), was carried out with school children in the fifth grade. Reading (Chapman-Cook Speed Test) shows the greatest correlation with intelligence. The gain score correlation reported is .56, the highest one recorded so far. For the substitution and addition tests, the correlation of the gain score is slightly higher than the one obtained with the final scores. All the correlations for the gain scores were found to be positive except for the cancellation test. Taking the ratio of improvability of the ten highest scoring pupils and ten lowest scoring pupils, De Weerdt found that it is definitely higher for the more intelligent group, except the ratio for cancellation. Criticizing the inadequacies of these correlations, this is what he writes:

> A very few individuals within a group who achieve large increments in comparison with their original low initial scores or those initially high who achieve small increments because of being near their physiological limit exert a disproportionate influence upon the results in the use of the correlation formula. With the additional factor of favorable and unfavorable attitudes toward certain tests which are like tasks previously met, the correlation formula as means of stating relationship incorporates more than the simple relationship between the actual content of the tests involved... It is evident then that the general test of intelligence does indicate capacity for improvement but the general test does not indicate how much improvement we may expect in a specific function. This at once suggests the practicability of measuring the capacity for improvement directly in the specific function rather than attacking it by inference through a general test (13, p. 557).

One of the last studies in our summary table is the one carried out by Woodrow with fifth, sixth and seventh graders. He, more than

any other psychologist, studied extensively the problem of learning and its relation to intelligence with the aid of factor analysis. He no doubt gave good evidence that to identify intelligence with ability to learn as measured by our actual tests was quite erroneous. However to the writer's point of view, he does not seem to have gone far enough in the development of his problem to insure complete and satisfactory treatment. It would be sufficient to comment here on Woodrow's findings with the grade school children. In this study as in many others, he found that there is a "lack of significant relation between change in score and I.Q. in grades beyond the fifth*" (40, p. 153).

The writer computed the averages of the correlations in the three grades tested in this research and found the following:

Correlations between gain scores and intelligence

Fifth Grade	.402
Sixth Grade	.190
Seventh Grade	.100

Apparently from the above results we can infer that the older the child gets, the less he uses his intelligence in class achievement. It can be assumed that by the time a child gets to the sixth and seventh grades he gets new distractions that do not exist in the lower grades. Girls might become interested in the way they look and boys would probably want to become football stars or train engineers, and this may cause such a big drop from .402 to .100. Woodrow does not elaborate on this difference but mentions it in passing in the following paragraph:

---

*Italics are the present writer's

> It should of course be kept in mind that the apparent determining conditions of gains in school, that is, school attendance for a stated time in a given school class, although a more natural or life-like condition than practice conducted in a laboratory is in reality, a complex of variables which are not kept constant for all pupils. Some variables which readily come to mind are irregularities in attendance at school, interest and liking for the subject, health, study habits, level of aspiration, various home influences and a complex of conditions not too well understood, determining what is rather vaguely called motivation (40, p. 156).

Although the next experiment which will be reviewed did not deal in gain scores, it would be interesting to mention it here on account of the contribution it can make to our problem. Burtt, Chassell and Hatch (11), three psychology instructors at Ohio State University, arranged to teach elementary psychology to classes selected on the basis of intelligence. Each instructor taught a class of high, medium or low intelligence and one or two control classes of heterogeneous intelligence. Correlations were calculated between intelligence and grades after the first semester and second semester as well. However, during the second semester a new variable was introduced for the high and medium intelligence classes. They were all "pushed", that is, they were given longer assignments than usual. It was found that the most intelligent finished the course sooner. The low-intelligence class maintained normal rate of progress. The theory implied by the authors in their investigation is that if each student is stimulated to do his best the correlation between intelligence and academic marks could be raised.

Table II shows the results of the change in speed of instruction. As we can see the high intelligence group and medium one raised their correlation with intelligence from .32 to .73 and from .15 to .75 respectively while the controls did not show much change. The group

of low intelligence on the other hand had their correlation drop from .67 to .22. The authors give various reasons for this fall, one being that during the first semester, the students were forced to work up to maximum intellectual capacity.

Commenting on these results, this is what the authors have to say:

> The obvious conclusion is that from the pedagogical point of view nothing is gained by grouping students of superior intelligence for instructions under the conditions of this experiment unless they are forced to cover the material of the course at a more rapid pace than average. Not only are they obviously able in the latter case to cover ground more rapidly but the individual student comes near working at his maximum intellectual efficiency (11, p. 161).

TABLE II

CHANGES OCCURRING IN CORRELATIONS BETWEEN GRADES AND INTELLIGENCE WHEN RATE OF PROGRESS OF COURSE IS CHANGED

(From Burtt, Chassell and Hatch) (11)

Groups	First Semester	Second Semester
Selected high intelligence	.32	.73
Control for high group	.46	.45
Selected medium intelligence	.15	.75
Control for medium group	.38	.15
Selected low intelligence	.67	.22
Control for low group	.64	.58

STATEMENT OF THE PROBLEM

Now what does this array of research tell us? One thing that is quite clear is that we do not have uniformity with these studies. Some achievements correlate high or low with intelligence initially and similarly at the final test. In some cases the correlation goes up and in others it goes down. Furthermore correlations with the gain scores remain relatively high when the correlation with achievement improves in the final test. In other instances, it is nil, although both the correlations with achievement remain unchanged but high. All these differences naturally need explanation. Although we have found in most studies that intelligence is not the ability to learn, it would not be wise to consider the matter settled since in certain instances we have obtained opposite evidence. The question arises, therefore, whether the same results would be obtained if different kinds of tests were used, for example, reasoning in psychology.

To answer these numerous questions, the present study has been set up by giving separate tests of different aspects of the same course in order to see to what extent this differentiation affects the correlations with intelligence.

SOURCES OF DATA USED IN THIS RESEARCH

The data of the present students were obtained in part from an investigation carried out by Smith and Johnson (unpublished monograph) * of the Department of Psychology at Michigan State College. The purpose of their study was primarily to evaluate the effect of democratic teaching procedure on students' attitudes and achievements.

One hundred and sixty-two students, mostly sophomores, were given an achievement test during the first two periods of the term. Then they were divided and matched according to the experimental design required by such a research. At the end of the course, the students were again tested with an equivalent test of psychology. There were therefore pre-test and post-test scores in elementary psychology available for the purpose of this research. The 162 students were divided into four classes, and each instructor taught two classes using the democratic procedure in one and the lecture method for the other. Since these different methods of teaching produced only small effects upon the scores, the gains being approximately equal for the four groups, the data available can therefore be used as coming from a single universe or population.

As pointed out earlier in this paper, research has indicated that gain in scores usually does not correlate significantly with intelligence. The novelty of the present study is that different kinds of achievement tests were used. In this way, correlations between intelligence and

* Will be published in Psych. Mono. 1953.

various kinds of achievements were obtained. It was pointed out in the introduction that there was no attempt in the studies made to explain differences obtained with different types of tests, nor why some gains correlated higher with intelligence than others.

Description of the Tests

Out of a large pool of test items devised by Smith and Johnson, all multiple-choice items with five alternatives, three different tests were constructed. Each was made up of sixty items covering respectively: 1) Vocabulary, 2) Factual knowledge, 3) Reasoning in psychological matters. These three tests were divided in two forms of approximately equal difficulty. Item analysis was carried out on the basis of scores obtained previously from a similar population of students. Thus one form was used as the pre-test, and the equivalent form was used as the post-test.

Vocabulary. The following is an example of the vocabulary items:

SET 1) preparation 2) reaction 3) pathways 4) choice 5) fixation.

From Table III, we can see that this test when initially given had a reliability of .46, which jumped to .76 on the final test. The odd-even method was used, corrected for length by the Spearman-Brown formula. Comparing the two reliabilities for the two forms, we can see that the post-tests show significantly greater reliabilities than the pre-tests. This naturally should be expected on account of the fact that many responses in the pre-test must have been chance responses, since students had to rely upon inspiration based on their scanty contact with psychology in the past.

Facts. The questions on factual knowledge depended largely on memory of specific facts which appeared in the textbooks. As an example, we can quote one question:

> Raising body temperature by electric current is a specific treatment for (1) Psychosomatic illness (2) Psychoneurosis (3) Brain syphilis (4) Manic-depressive (5) Paranoid schizophrenia.

As shown by Table III, the reliability of the pre-test was .51 while it became .68 for the post-test.

Reasoning. The third type of sub-test was made up of items which required applications of learned principles to new problems. Naturally a certain amount of vocabulary and factual knowledge are assumed to be necessary to answer these questions. An example of the reasoning item is as follows:

> Under which of these conditions will the sex drive in rats be stronger than hunger or thirst? (1) Give plenty of food and water (2) Place animal in obstruction box when young (3) Deprive animal of food, water and sex for two days (4) Deprive animal of food and water for one day (5) Keep male and female together constantly.

This particular test showed the greatest increase in reliability from the pre-test to the post-test, being .33 and .70 respectively.

The American Council of Education Psychological Examination. The initial and final test scores were correlated with the total scores on the A.C.E. and with the linguistic part of the A.C.E. One hundred students' intelligence scores (83 sophomores, 11 freshmen, 5 juniors and 2 seniors) were used in this study out of the 162 of the original study. This was the number of A.C.E. raw scores that were obtained from the Examiner's Office of Michigan State College. Since the data of these

TABLE III

COEFFICIENTS OF RELIABILITY OF INITIAL AND FINAL TESTS N = 162

(from Smith & Johnson)

Sub-Tests	Pre-Test r	Post-Test r
Vocabulary	.46	.76
Facts	.51	.68
Reasoning	.33	.70

tests are printed in booklet forms showing decile rankings, the writer had to refer to the original records in order to obtain original scores suitable for correlational purposes.

Although the A.C.E. examinations were taken by the students a year or so before the actual experiment for the majority of them, the use of the same scores is justified on account of the well-known reliabilities, .95 for the linguistic score and .97 for the total score.

The reliabilities do not seem to change much after some time. Livesay (22, p. 67), after giving the same tests to seniors who had taken it four years earlier, found that the reliability was still as high as .88 for the total test.

Regarding the validity of the A.C.E. scores, Berdie and others (8), have found that the range of correlations with grade-point average is from .25 to .66. Furthermore Garrett reports that the A.C.E. correlated more closely with college average than did the scores of other intelligence tests (17, p. 129). Reverting to Berdie's survey covering thirteen colleges, universities and teachers' colleges, Michigan State College being one of them, it was computed by the writer that the average score obtained for the linguistic part of the A.C.E. was .67.24. The average obtained for the present sample of one hundred M.S.C. students was rather close, 66.79.

In this study the total scores on the A.C.E. were correlated with the pre-tests and post-tests, and then the correlations were computed with the linguistic scores only. It was found, however, that the latter

could be used independently. Berdie and others point out in this connection:

> In a surprising number of cases the L scores yield almost as high or higher correlation than the total score. Even in biological sciences, wherein a superficial analysis would suggest a nice balance of quantitative and linguistic abilities the total scores predict little if any better than the linguistic scores (8, p. 811).

RESULTS

Statistical Techniques. There were for this study twelve sets of data. Four sets comprised the pre-tests, which included the total scores. Four other sets were made up of the post-tests. The last four sets were calculated by subtracting the former from the latter in order to obtain the gain scores for each sub-test and total. The pre-tests and post-tests were all intercorrelated and correlated with the A.C.E. linguistic scores as well as with the total scores. Gain scores were correlated only with the linguistic scores. There were in all twenty-eight correlations. To avoid any possibility of errors, every one was run twice.

Gains in general. Before turning to an analysis of the data relative to the main problem of this research, it would be appropriate to examine some general aspects of the results. Table IV gives the mean and standard deviation of each test, as well as the critical ratios between the initial and final tests. We can see that the vocabulary shows the greatest amount of improvement while reasoning shows the least, although quite significant. It can be inferred from these results that the extent of improvement as indicated by the critical ratios reflect the technical degree of the subject matter. The course in psychology has the greater effect on the mastery of technical words than on the application of the concepts subsumed in their definitions. However, one criticism that could be levelled at the inference drawn above is

TABLE IV

MEANS, STANDARD DEVIATIONS, DIFFERENCES BETWEEN MEANS AND CRITICAL RATIOS OF PRE-TESTS AND POST-TESTS
N = 100

Sub-Tests	Mean of Pre-Test	S.D.	Mean of Post-Test	S.D.	D.	C.R.
Vocabulary	13.31	3.11	22.72	3.90	9.41	18.72
Facts	11.02	3.25	19.37	3.96	8.35	16.29
Reasoning	15.66	2.97	20.83	3.43	5.27	11.71
Total	39.89	6.83	62.92	8.20	28.03	20.10

that the critical ratios may not have any comparative value since the pre-tests are not all of equal difficulty. But still it is safe to assume that facts could be learned and greater improvement can be made than could possibly be done with reasoning tests. Learning how to reason is very problematic, while there is no doubt about learning facts.

The distributions of initial and final scores are represented in Figs. 1, 2 and 3, and for the total, Fig. 4. It can easily be seen that there is a greater overlapping in the reasoning tests than in the other sub-tests.

Correlation between initial and final scores. Correlations computed between the initial and final test scores are shown in Table V. The coefficient of correlation between the pre-test and post-test in reasoning is the highest, .40. Learning, therefore, had the least effect in disturbing the rankings of students in this sort of achievement. If .40 is corrected for attenuation, it is found that the estimated correlation is .83, which is obviously very high. Facts hold second place with a correlation of .33, and lastly, vocabulary with a correlation of .14, this being insignificant at the 5% level. We can interpret this disparity, as already pointed out, to the fact that the study of psychology has the greater effect upon the knowledge of psychological terms, which does not necessarily imply a thorough understanding of the concepts. In the framework of the actual study, the learning of vocabulary and facts depended more upon motivation or non-intellectual factors than upon intelligence. Reasoning depends more on intelligence both initially and finally.

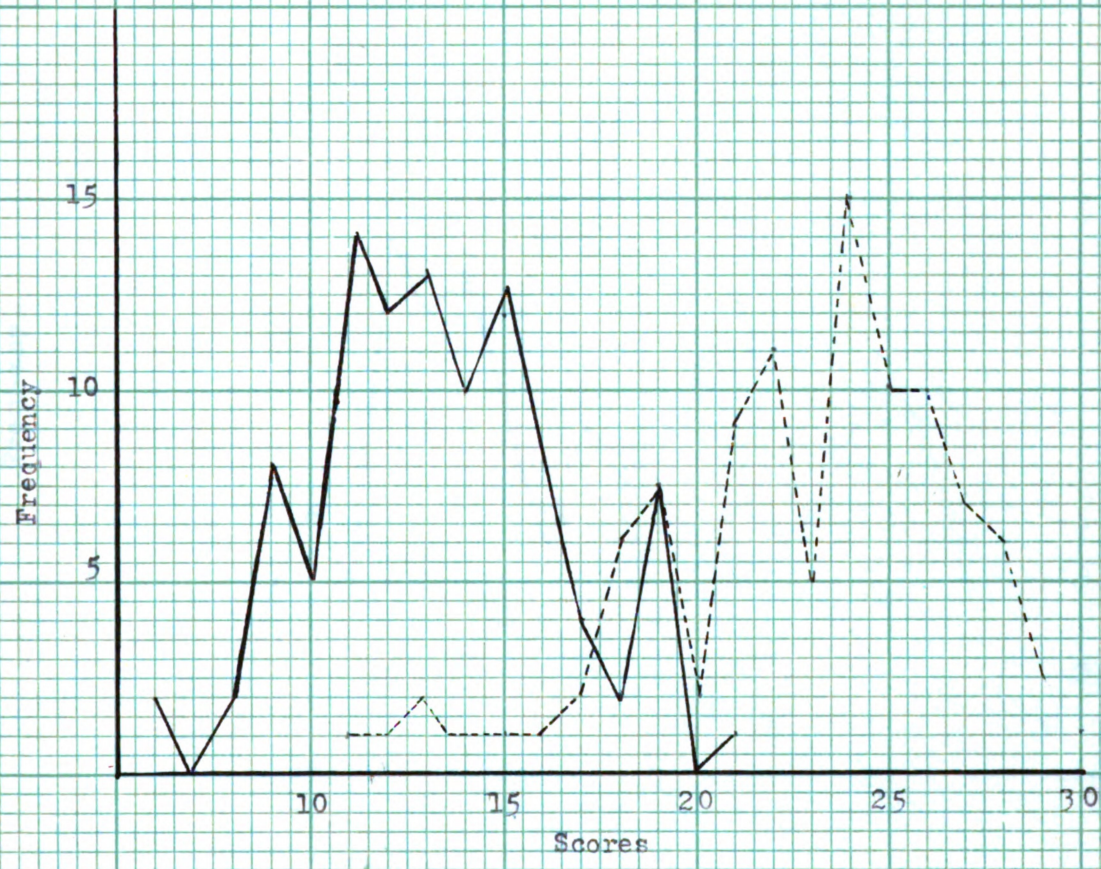

Fig. I. Graph showing achievement on initial and final sub-tests in Vocabulary.

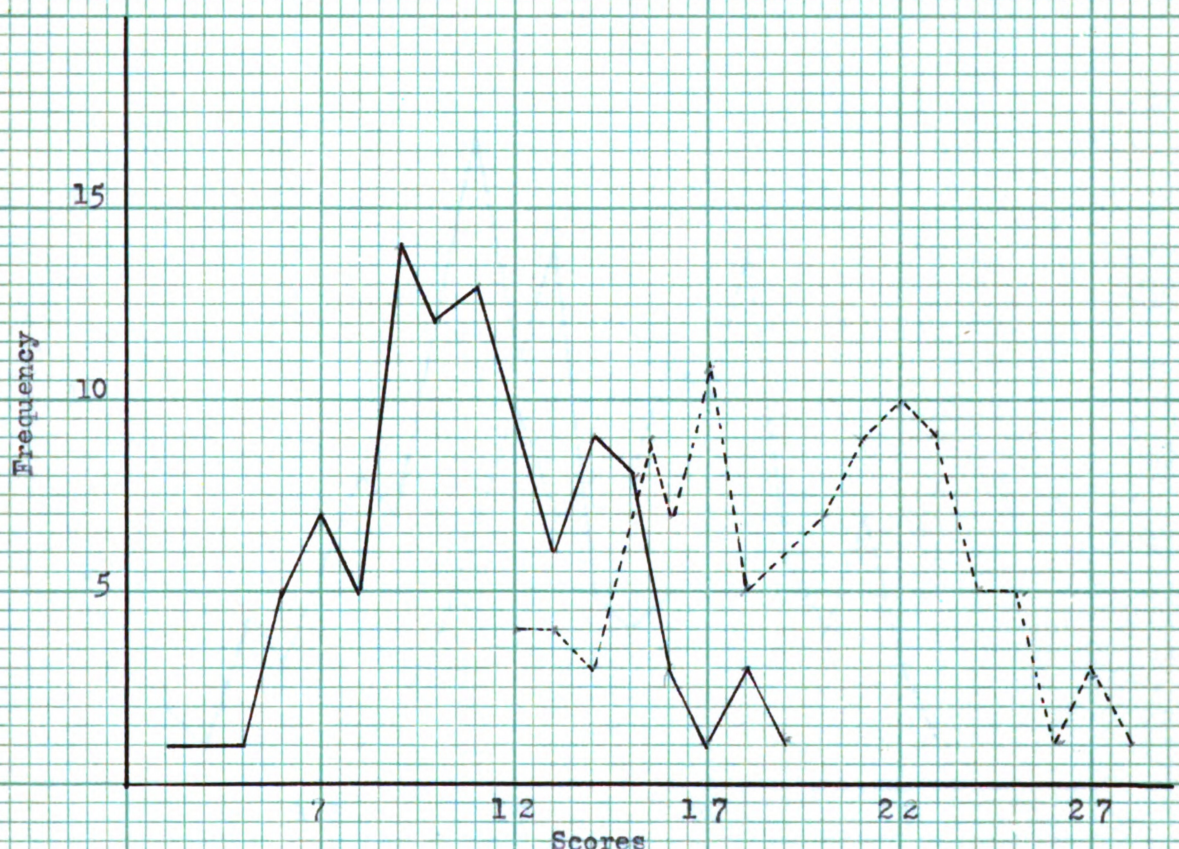

Fig. 2. Graph showing achievement on initial and final sub-tests in facts.

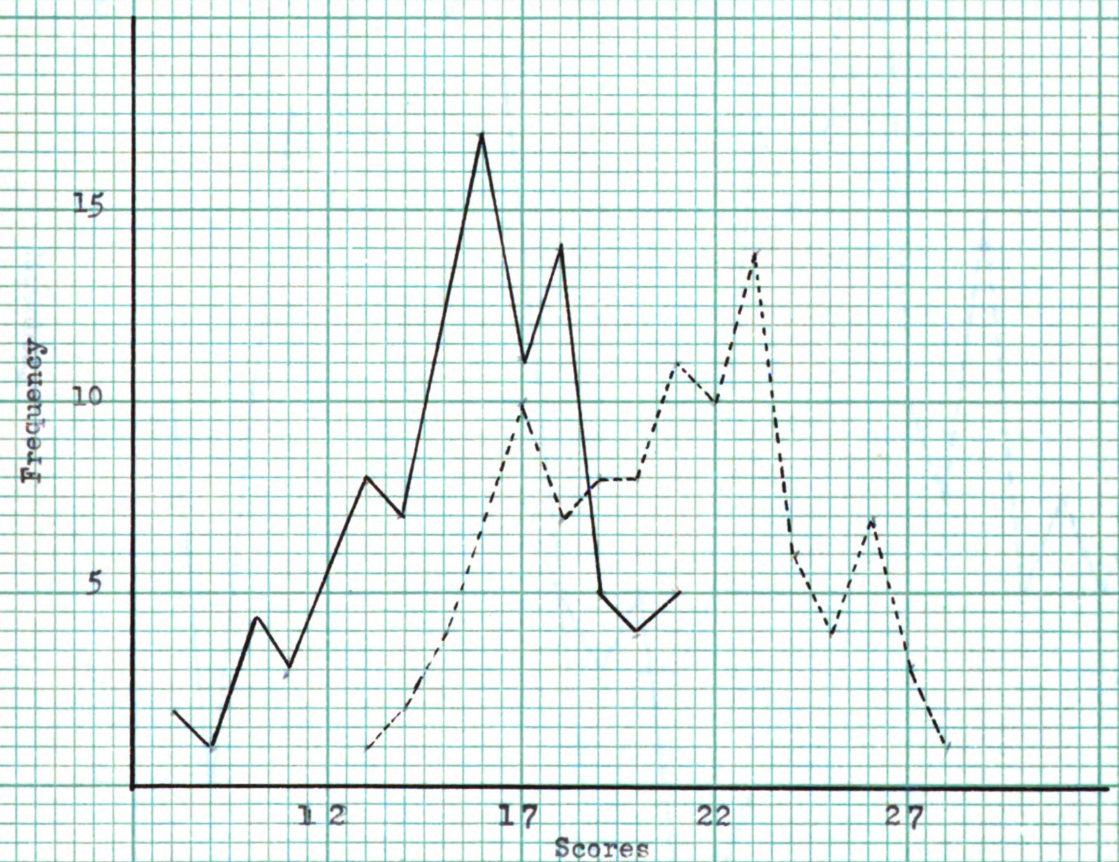

Fig. 3. Graph showing achievement on initial and final sub-tests in reasoning.

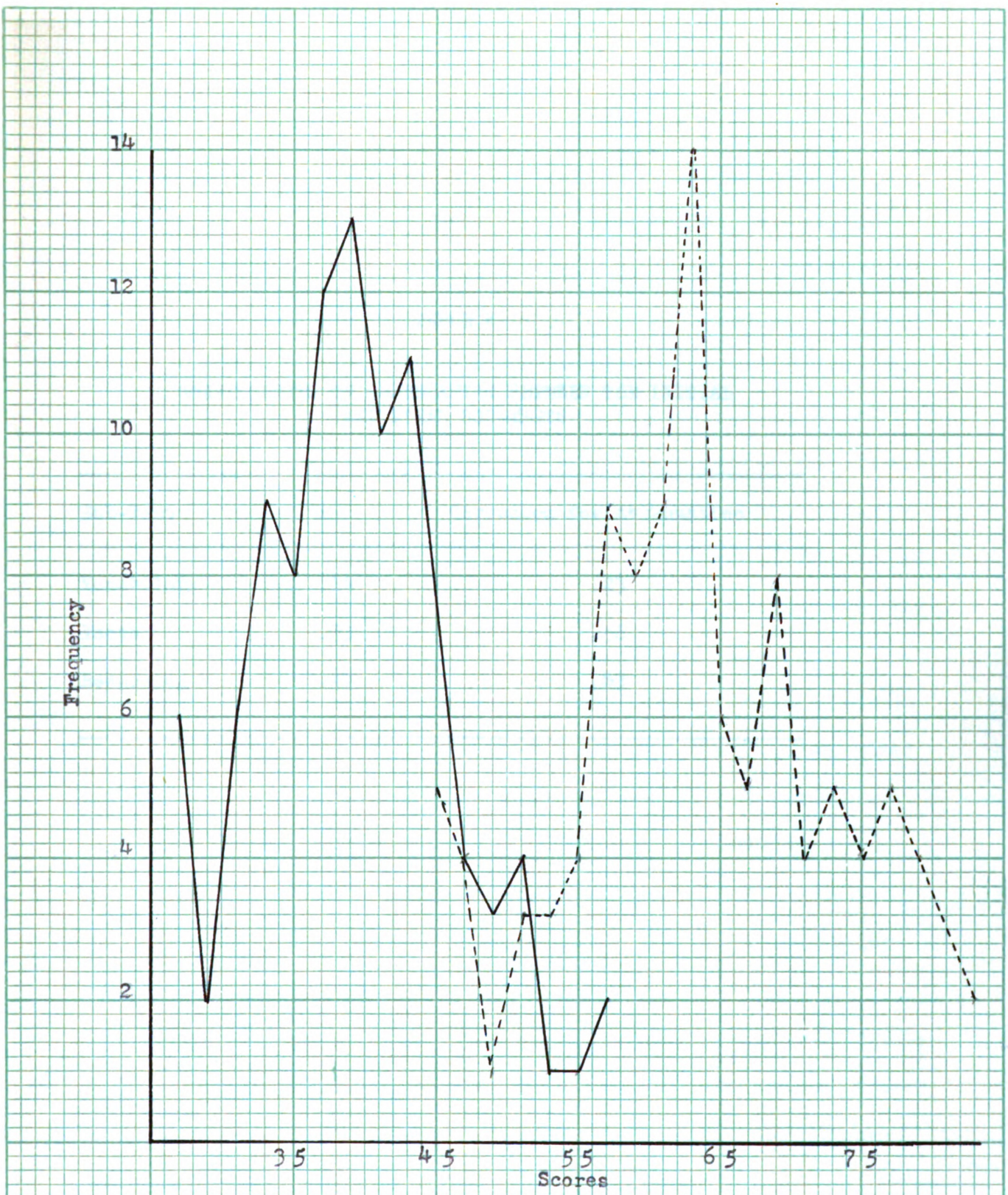

Fig. 4. Graph showing achievement on initial and final total tests in Psychology.

TABLE V

COEFFICIENTS OF CORRELATION BETWEEN THE INITIAL AND FINAL SUB-TESTS

Sub-Test	r	Corrected for Attenuation
Vocabulary	.14	.23
Facts	.33	.56
Reasoning	.40	.83

Comparing the data given for the three sub-tests in Table III and V discloses an interesting relationship. The tests with the highest test-retest gain have the lowest test-retest correlation. The test of vocabulary shows the largest gain and the lowest correlation. Reasoning shows the smallest gain and the largest retest correlation. Going farther in the same direction, one might add that the retest studies of the A.C.E. show an even smaller gain and higher correlation. In these respects psychological reasoning lies somewhere between an intelligence test and conventional achievement test. This is consistent with the above statement of the greater involvement of non-intellectual factors in the learning of vocabulary and facts than in learning to reason.

All the above correlations were corrected for attenuation, using the reliabilities given by the original study with 162 students, while all our computations were based on the data obtained from one hundred students. The use of the original reliabilities is justified, since the intercorrelations, means and standard deviations found by Smith and Johnson for 162 students are very close to the ones found by the writer for one hundred of these 162.

Intercorrelation between the sub-tests. Although the intercorrelations were computed in the original study with 162 students, these were again computed with our sample of one hundred students. Results are shown in Table VI. Just as the odd-even reliabilities of the post-tests were found to be higher (see section source of data) as per Table III, a similar tendency was found for the intercorrelations of the post-tests

TABLE VI

INTERCORRELATIONS OF INITIAL TESTS N = 100

(Figures below the diagonals are corrected for attenuation)

a) Initial Tests	Vocabulary	Facts	Reasoning
Vocabulary	—	.32	.18
Facts	.65	—	.39
Reasoning	.47	.95	—

INTERCORRELATIONS OF THE FINAL TESTS

b) Final Tests	Vocabulary	Facts	Reasoning
Vocabulary	—	.50	.38
Facts	.70	—	.59
Reasoning	.52	.85	—

for the same reasons as pointed above, the chance factors being reduced in the final tests. However, in correcting these intercorrelations for attenuation, the figures obtained are much closer. The intercorrelations are in part a function of the reliabilities of the two correlated tests.

Of all these intercorrelations, the lowest, initially and finally, are those between Reasoning and Vocabulary, a further indication that they require different factors.

Correlations with the A.C.E. Psychological Examination. Table VII shows the correlations of the sub-tests with the total as well as with the linguistic scores of the A.C.E. Out of the six correlations, we can see that we obtain in three cases slightly higher correlations with the linguistic A.C.E. scores, in two cases exactly the same correlation. The final factual test was the only one which showed a slight increase in using the total score (from .08 to .11), probably due to the fact that there may be more quantitative elements in it than in vocabulary and reasoning which are more verbal in content. In view of the comments made by Berdie and others (8), and the actual findings in this study, it was decided that the linguistic scores would be used for the rest of the computations.

From the results obtained as shown in Table VII, the vocabulary test seems to be least related to intelligence which is in contradiction to the well-known fact that vocabulary knowledge correlates highest with intelligence. As Terman points out, "We have found that the vocabulary test to be the most valuable single test in the scale" (6, p. 302),

TABLE VII

CORRELATIONS BETWEEN THE THREE SUB-TESTS AND THE A.C.E. TOTAL AND LINGUISTIC SCORES

a) Initial Test	r with L scores	Corrected for attenuation	r with total scores	Corrected for attenuation
Vocabulary	.16	.24	.10	.15
Facts	.19	.27	.19	.27
Reasoning	.22	.39	.22	.39
b) Final Test				
Vocabulary	.10	.12	.05	.06
Facts	.08	.10	.11	.13
Reasoning	.27	.32	.21	.25

also "Our statistics show that in a large majority of cases, the vocabulary test alone will give an intelligence quotient within 10% of that secured by the entire scale" (6, p. 230).

One explanation for the low correlations of .16 and .10 between vocabulary and the linguistic A.C.E. scores may be the fact that the vocabulary tested here was of a special type which depended more upon the different interests of the students, etc. than upon their intelligence.

Another finding which may be significant and which might possibly throw some light to the whole problem studied is the fact that the correlations between the A.C.E. and both the Vocabulary and Facts sub-tests decreased in the final tests (Vocabulary from .16 to .10, Facts from .19 to .08), while the correlation of the Reasoning test on the other hand increased in the final examination (from .22 to .27). This sort of finding was already reported in the previous research mentioned in the introduction. A reason must be sought for these correlations, some of which are raised on the final test, while others are lowered.

In the case of vocabulary and facts, although mostly based upon the students' interest, there is still some dependence on intelligence. However, this is reduced in the final test on account of the learning period. Same results were found by Carlson, Fisher and Young (12).

Furthermore Woodrow obtained the same trend with different types of learning tasks, such as Horizontal Adding, Analogies, Letter-digit etc., under laboratory conditions. His conclusion was:

> "...the effect of practice may be to lessen the correlation between the test practiced and such intelligence tests as

> those included in the Otis battery, and indeed that a slight lowering of the correlation of the practiced test with intelligence, reliability kept constant, is a rather usual result (36, p. 572).

However, he states that this holds true and applies to correlations of two equally reliable measures. Our vocabulary and facts subtests do not actually fulfill these conditions since reliabilities in our study differ from initial to final tests (Vocabulary from .46 to .76, facts from .32 to .50), and yet they are still in accordance with the conclusion drawn by Woodrow. When corrected for attenuation the difference in correlation is more pronounced. (Vocabulary from .24 to .12, Facts from .27 to .10).

In the case of Reasoning, as already pointed above, the change follows the opposite trend, that is the correlation with the A.C.E. is higher in the post-test than in the pre-test. The final reasoning test involved a greater degree of intellectual factors than in either of the first two tests. In other words, in the case of vocabulary and facts, a student of a certain intelligence may easily be overtaken in his final score by a less intelligent but more motivated student, while in the case of the reasoning test, it is likely that the student may have either kept his ranking as before or more likely he may have improved his score on account of the use of both his intelligence and the little or great amount of material he acquired during the term. Actually the reasoning tests tap more of the intellectual factor than memory which is related to a closer degree with vocabulary and facts.

Taking the correlation of the whole initial and final tests with the A.C.E. Linguistic scores, we find that the r for the pre-test with intelligence is .26, which is significant at the 1% level while the r for the post-test is .17, which falls a little short of the 5% level of confidence.

The magnitude of the differences, it is true, is not large. The usual test of significance of differences in correlations does not apply here since the purpose of the latter is to find out whether two samplings belong to the same population, the test remaining constant. As already calculated there is a significant difference in the means, but whether the difference between a correlation of .19 and one of .08 is significant is hard to say. We can only state that there seems to be a tendency of the correlations to drop in certain cases. Woodrow writes in connection with similar findings:

> It is true that the decrease in correlation was often small. It would be remembered in this connection, however that we are dealing here with a change in correlation in the case of a fixed group of subjects. The ordinary criterion of the significance of a difference between two coefficient of correlations does not, therefore, apply in the present instance (42, p. 571).

Another quotation referring to this deficiency is the one taken from Lindquist's book of Statistics for Educational Research:

> The Mathematical statisticians have not yet devised a test of significance of a difference between r coefficient for this situation. This is particularly unfortunate since it is just in this situation (evaluation of test material) that a test of significance is mostly needed (4, p. 218).

Correlation of the Gain scores with the A.C.E. From Table VIII, we can see that the coefficients of correlation between the gain scores

TABLE VIII

CORRELATIONS BETWEEN INTELLIGENCE AND GAIN SCORES IN PSYCHOLOGY

Type of Test	M.S.C. Study with A.C.E.	Carlson with Otis	Dysinger with Army Alpha
Vocabulary	-.026		
Facts	-.063		
Reasoning	.080		
Total	-.024	-.039	-.064

and the intelligence test are negative for both vocabulary and facts but not significant (-.024 and -.063). The correlation obtained for reasoning was found to be positive (.08) but still far from being significant. Like many other findings described in the introduction, our results also show that to describe intelligence as the ability to learn is of very dubious value. When compared to Carlson's correlation of -.039 and Dysinger's of -.064, both obtained with Elementary Psychology items, we can see that the results are not different except possibly for reasoning.

One of the statistical difficulties to be taken in consideration is the fact that the gross improvement scores are not necessarily accurate measures of absolute progress since it is more difficult for the beginning student with a high score to improve than it is for the student beginning with a low score. An improvement of five points in the upper quarter is not equal to an improvement of five points in the lower quarter. The first student has less room to improve. Studies reviewed in connection with gain scores seem to ignore this fact and it is only Dysinger (15) who considers it. To obviate such discrepancies in attainment, this experimenter suggested the use of an improvement ratio as a measure of relative improvement of each student. This is the ratio between the gross improvement and the maximum possible improvement. As an example, the student who obtains a score of 20 out of 30 possible answers in the pre-test and a score of 25 in the post-test will have an improvement ratio of 5:10 or 50. However, Dysinger found that both methods of computation of the coefficient of correlation led to the same

conclusion. Gross improvement or improvement ratio did not correlate significantly with intelligence.

Reasoning does not follow the same trend as the other sub-tests since we obtain a slight positive correlation differing from vocabulary by .10 and from facts by .14. Whether these differences have any significance cannot be computed. Lacking a good measure of the significance of the difference between correlated r's, we can only conclude that the difference is small and probably not significant.

DISCUSSION

If the statement is made that intelligence is not the ability to learn, then all correlations between gain scores and intelligence must be negligible. But, as already pointed out, the correlations are positive in certain cases. Since this difference in findings needs explanation, this paper will endeavour to posit certain hypotheses to account for it.

In the first place, the reason why intelligence is identified with ability to learn is the fact that achievement correlates with intelligence, and it is therefore very easy to fall into the error of assuming that learning can be identified with intelligence. Woodrow's (41) mathematical formulation shows that although intelligence may correlate with initial and final tests, it need not correlate with the gain scores.

Whatever intellectual factors contribute to the final score also contribute to the initial score and are subtracted out to obtain the gain score. The only way any factor can correlate with gain score is by contributing something additional between the initial and final score. The results of the present study show the possible but not significant contribution of reasoning in this way. Reasoning may make a larger contribution to the final tests than to the initial.

Non-intellectual factors operating between the initial test and the final test should also be considered. When a student takes an achievement test, it is presumed that during that hour or two hours, he will make

maximal use of his intelligence, motivation, persistence, etc., and what not. The gain score, however, is a function of a rather extended period of time and thus many uncontrollable variables enter into play which are non-existent at the time of the single hour test. Therefore the correlation between gain scores and intelligence does not exactly indicate the full relation between the ability to learn and intelligence. Among the non-intellectual factors that may be important, one should mention May's (23) study concerning the relation of time of study with scholarship. He found that the correlation was .32 and that if this factor could be kept constant the partial correlation between intelligence scores and honorpoints would be .805.

According to Ryans, learning seems to be conditioned by (a) the stimulus situation and (b) individual aptitude, persistence and motivation (29, p. 69). Except for persistence, which has received a scant treatment, all the other concepts are rather well understood in the field of psychology. According to Ryans, persistence has approximately the same meaning as endurance or continued energy release. This is what he says in connection with the interaction between intelligence and persistence:

> The two capacities interact and function together in determining response to stimulation. Aptitude, in the sense that it has been used here and in the sense that it is usually applied, defines the limits of learning in any particular sphere of activity. It determines what an individual can do or is capable of doing. Persistence on the other hand, refers to the capacity of an individual for continuous response and it determines roughly the extent to which he will exert himself in acquiring a response (29, p. 72).

Giving a plain example, a swimmer to become a champion should have both aptitude and persistence. A student of average intelligence may excell a brighter student because of his greater capacity for hard work. It may be possible that a student may persist in one field while exerting no effort in another and this depends on his motivation or the incentive available.

In an attempt to predict scholarship better than it has been done so far through intelligence testing, Ryans (31) studied a number of test situations which might be indicators of persistence. After giving 13 of these tests to 40 of his students and subjecting the data to multiple-factor analysis, he was able to obtain evidence of a general factor of persistence which contributed to many of the measures he employed. He furthermore found another factor which seemed to be heavily weighted with intelligence which was entirely unrelated to persistence. His aim was to develop an instrument which would at least estimate the degree of persistence possessed by an individual. Boiling down his tests to four, Anagrams, study time, endurance and study log, he obtained a composite which correlated with scholarship .40 while its correlation was nil with intelligence. With an r of .48 between intelligence and honor-point ratio, he obtained a multiple r between school success and persistence and intelligence combined of .66. In another study, he was able to obtain a multiple r of .73 and .79 (33).

Howell previous to Ryans has the same purpose in mind. He figured out that since there was a low correlation between persistence and intelligence and relatively high correlations between each of these measures

with grades, this suggested to him the possibility that a combination of persistence and intelligence scores might afford an improved prediction of grades. He confirmed his theory when testing 24 students who were failures in college although having high I.Q. Their average on the persistence tests was lower than the average. Like Ryans, the multiple coefficient of correlation he found was .64. A more recent study by John French (16) of the Educational Testing Service reports a multiple correlation of .65 for the prediction of grades by combining intelligence and persistence tests. With the results obtained in this research and the ones described early in the paper and the studies reviewed concerning the importance of other factors in achievement, we can now present a few hypotheses:

(1) Those gain scores that have shown a certain degree of correlation with intelligence cover those subjects that are more connected with intellectual factors. For instance, intelligence probably contributes more to reasoning tests than to others.

(2) Those gain scores which have been found not related to intelligence would probably be related to some other factor. Since we have learned that acquisition is a function of many other factors besides intellect, we could assume that if these gain scores were correlated with motivation, persistence etc., we might obtain significant correlations. It must be added there that the tests might have varying degrees of each factor involved. This probably explains the range of correlations obtained in studies reviewed.

(3) A further broad generalization that could be made from these results is that whenever the post-test correlated higher with intelligence than the pre-test, the gain scores would correlate positively with intelligence. If on the other hand the correlation of intelligence with the post-test is lower, the factor of persistence or motivation operates in this drop.

SUMMARY AND CONCLUSIONS

Because of the insignificant results we have obtained in correlating gain scores with intelligence, we can say with certain reservations that intelligence is not entirely the ability to learn. While the pre-test and post-test might correlate positively with intelligence, the gain scores in most instances show little relation with general intelligence.

However, the following statements could be made resulting from the actual research:

1. Scores and gain scores in reasoning correlate higher with intelligence than scores and gain scores in vocabulary and facts. The latter might correlate higher with motivation, persistence, etc.
2. The use of sub-tests yields more information about correlations of intelligence with learning than a single test of overall achievement.
3. The findings have implications in the field of vocational and educational guidance. They suggest a limitation on the use of intelligence tests and the importance of non-intellectual factors in the prediction of improvement in college. Work on persistence and motivation, which are instances of non-intellectual factors, by Howells, Ryan and French has promising possibilities in this direction. Prediction and recommendations made on the basis of an intelligence test score should be limited to the degree to which the learning of a task involves intellectual power.

BIBLIOGRAPHY

A. Books

1. Guilford, J. P., Fundamental Statistics in Psychology and Education. New York, McGraw-Hill Book Co., Inc., 1950. 629 pp.
2. Kelley, T. L., Statistical Method. New York, Macmillan, 1938.
3. Lindquist, E. F., Educational Measurements. American Council of Education, Washington, D.C. 1951.
4. Lindquist, E. F., Statistical Analysis for Educational Research. Boston, Houghton Mifflin, 1940.
5. Stoddard, George D., The Meaning of Intelligence. New York, the Macmillan Company, 1943.
6. Terman, Lewis and Maud A. Merrill, Measuring Intelligence. Cambridge, Houghton Mifflin Co., 1937.
7. Woodworth, R. S., Psychology. New York, Henry Holt and Co., 1940.

B. Periodical Articles

8. Berdie, Ralph, Paul Dressel and Paul Kelso. "Relative Validity of the Q. and L. scores of the A.C.E. Psychological Examination." Ed. and Psychol. Measurements, Vol. 11, No. 4, 803-812, Winter 1951.
9. Blair, Glenn Myers. "How learning theory is related to curriculum organization." Journal of Educ. Psychology, 39: 161-166, 1948.
10. Bolton, Euri Belle. "The Relation of Memory to Intelligence" Journ. of Exp. Psychology, 14, Feb. 1931.

11. Burtt, E. Harold, Laura M. Chassel and Elizabeth Hatch. "Efficiency of instruction in unselected sections in elementary psychology compared with that in sections selected on the basis of intelligence quotient." Journ. of Educ. Psych., 14: 154-161, 1923.

12. Carlson, H. B., R. P. Fisher and D. T. Young. "Improvement in elementary psychology as related to intelligence." Psych. Bulletin, 42: 27-34, 1945.

13. De Weerdt, E. H. "A study of improvability of fifth grade school children in certain mental functions." Journ. of Educ. Psych., 18: 547-557, 1927.

14. Drake, C. S. "The iota function." Journ. of Educ. Psych., 34: 190-198, 1940.

15. Dysinger, W. D., and W. S. Gregory. "A preliminary study of some factors related to student achievement and grades in the beginning course in psychology." Journ. of Gen. Psych., 24: 195-209, 1941.

16. French, John W. "The validity of a persistence test." Psychometrika, Vol. 13, No. 4: 271-277, December, 1948.

17. Garrett, H. F. "Review and interpretation of investigation of factors related to scholastic success in college." Journ. of Exp. Educ., 18: 91-138, December, 1949.

18. Heese, K. W. "A general factor in improvement with practice." Psychometrika, 7: 213-223, 1942.

19. Howells, T. H., "An experiment of persistence," Journ. of Abn. and Soc. Psych., 28: 14-23, 1933.

20. Husband, R. C., "Intercorrelations among learned abilities," Journ. of Gen. Psych., 55: 353-364, 1939.

21. Johnson, O. J., "A study of the relation between ability to learn and intelligence as measured by tests," Journ. of Educ. Psych., 14: 540-541, 1923.

22. Livesay, T. M., "Does test intelligence increase at college level," Journ. of Educ. Psych., 30: 63-68, 1939.

23. May, Mark, "Predicting academic success," Journ. of Educ. Psych., 14: 429-440, 1923.

24. MacConnell, T. R., "Changes in score on the psychological examination of the A.C.E. on education from freshman to senior years," Journ. of Educ. Psych., 25: 66-69, 1934.

25. MacCullough, Constance, "Relationship between intelligence and gains in reading ability," Journ. of Educ. Psych., 30: 688-692, 1939.

26. Shuey, Audrey, "Improvement in scores on the American Council Psychological Examination from freshman to senior year," Journ. of Educ. Psych., 39: 417-426, 1948.

27. Simral, Dorothy, "Intelligence and the ability to learn," Journ. of Psych., 23: 27-43, 1947.

28. Spence, R., "Factors related to college achievement," Teachers College Record, 29: 504-519, 1927.

29. Ryans, David G., "The major observable dimensions of behavior," Journ. of Gen. Psych., 19: 65-77, 1938.

30. _____, "The meaning of persistence," Journ. of Gen. Psych., 19: 79-96, 1938.

31. _____, "An experimental attempt to analyze persistent behavior: I. Measuring traits presumed to involve "persistence," Journ. of Gen. Psych., 19: 333-354, 1938.

32. _____, "An experimental attempt to analyze persistent behavior: II. A persistence test," Journ. of Gen. Psych., 19: 354-372, 1938.

33. _____, "A study of the observed relationship between test results, intelligence and academic success," Journ. of Educ. Psych., 29: 573-580, 1938.

34. Woodrow, H., "The effect of practice on groups of different initial ability," Journ. of Educ. Psych., 29: 268-278, 1938.

35. _____, "The effect of practice on tests intercorrelations," Journ. of Educ. Psych., 29: 561-572, 1938.

36. _____, "Factors in improvement with practice," Journ. of Psych., 7: 67-68, 1939.

37. _____ "Interrelations of measures of learning," Journ. of Psych., 10: 49-73, 1940.

38. _____, "The problem of general quantitative laws in psychology," Psych. Bull., 39: 9-27, 1942.

39. _____, "Intelligence and improvement in school subjects," Journ. of Educ. Psych., 36: 155-166, 1945.

40. _____, "The ability to learn," Psych. Bull., 1946.

41. _____, "The relation between abilities and improvement with practice," Journ. of Educ. Psych., 29: 215-230, 1938.