

1950

An analysis of academic achievement in psychology

Charles Spital

Follow this and additional works at: <http://scholarship.richmond.edu/masters-theses>

 Part of the [Psychology Commons](#)

Recommended Citation

Spital, Charles, "An analysis of academic achievement in psychology" (1950). *Master's Theses*. Paper 1043.

This Thesis is brought to you for free and open access by the Student Research at UR Scholarship Repository. It has been accepted for inclusion in Master's Theses by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.

AN ANALYSIS OF ACADEMIC ACHIEVEMENT IN PSYCHOLOGY

BY

CHARLES SPITAL

**A THESIS
SUBMITTED TO THE GRADUATE FACULTY
OF THE UNIVERSITY OF RICHMOND
IN CANDIDACY
FOR THE DEGREE OF
MASTER OF ARTS IN PSYCHOLOGY**

AUGUST, 1950

PREFACE

Some theses are begun with sudden and exciting inspiration. This one was not.

The need for a comprehensive examination for seniors majoring in psychology had been recognized by the faculty for several years. The writer had another idea which could have developed into a thesis, but upon hearing the above matter discussed by the psychology staff, the switch was soon made. Sensing that here was something which, if satisfactory, would be used, rather than gather dust on a library shelf; realizing how much psychology could be learned while devising a comprehensive examination in the field, I found the decision easy and the task challenging.

Credit should be given to the staff of the Psychology Department, Professors Merton Carver, Austin Grigg, and Stanley Skiff, for their assistance in the preparation of this thesis. I should like to express my appreciation also to the thirty-one students who cooperated in taking the

examination and in supplying other pertinent data.

C. S.

TABLE OF CONTENTS

INTRODUCTION	1
PART ONE	
Relationship Between Course Grades and Test Scores	8
PART TWO	
Order of Difficulty of Subjects Represented On Examination	12
PART THREE	
Students' Estimations of Their Test Performances .	19
CONCLUSION	22
APPENDICES	27
A. Chart of Students' Performances	27
B. Summary of Calculations	29
BIBLIOGRAPHY A	31
BIBLIOGRAPHY B	33
VITA	35

INTRODUCTION

Some kind of educational tests have been in use for centuries, dating back to the oral tests of Socrates. It was not until 1845, however, that the brilliant Horace Mann first pointed out the advantages of written essay examinations over oral. And even then he could not precipitate any great rush toward the use of written tests.

The innovation of written essay tests was not in itself satisfactory enough to warrant widespread use as a measuring instrument. For like the oral tests, they were still too subjective to really measure achievement. It was not until 1894 that J. M. Rice¹ finally realized the necessity of devising an objective achievement test. The inefficiency

¹Ross, C. C., Measurement in Today's Schools, p. 21.

of the old standards of measurement - oral and written essay tests - was apparent in that a pupil who did poorly in one class would do well in another; it was apparent in that there were no standards to serve as guides; and it was apparent in that there were no standards to settle the controversy over what subjects should be taught and the time needed for teaching them.

Making spelling the subject of his first experiment, Rice devised a list of fifty words, and toured the country's schools giving the standardized test to about 100,000 students. When he had analyzed the results, Rice found little relation between the ability to spell and time devoted to spelling in the school program. Spelling achievement was found to be just as good in schools devoting fifteen minutes to the subject as in other schools with a period of three or four times longer.

Next came the monumental work of Thorndike and his students, whose various tests and scales need little mention here. Then, others found the great discrepancy (now well verified) among teachers grading the same essay tests.² Reliability of grades, for teachers' evaluation of the same

²Falls, J. D. "Research in Secondary Education", Kentucky School Journal, 1928, 6:42-46

Starch, Daniel and Elliot, E. C., "Reliability of Grading Work in Mathematics", School Review, 1913, 21:254-259.

standardized English composition over a long period and several trials, was found by Hulten³ to be significantly low. He found that fifteen teachers who gave passing marks the first time failed the "pupil" the second time the paper was marked.

Many more facts were discovered which showed the need for more objective tests. Although there were, and still are, some objections to the disuse of essay tests, the main body of evidence points to the greater advantages derived from the use of objective examinations. Not only are validity and reliability greater for objective tests, but less time is consumed in taking and grading such tests, and more material may be covered in a shorter time.

A comprehensive examination, of the type presented here, retains all the advantages of an ordinary objective test and in addition has further advantages. One such advantage is the nineteen different scores which provide for analysis of the testee's performance on sixteen subjects and on the subjects he has taken and those he has not taken. Another asset is the wide coverage of four years' work in psychology, while most tests cover a month's study.

³Hulten, C. E., "The Personal Element in Teachers' Marks", J. of Educ. Research, 1925, 12:49-55.

As precedents for this examination, there are a number of noteworthy examples, among which are the Graduate Record Examination, the Law School Admission Test, the College Entrance Examinations, and the Medical School Admission Test. The latter test has already been well validated as to the relationships between test scores and medical school success.⁴

It is the writer's belief that the construction of an objective comprehensive examination covering the various courses offered in psychology at the University of Richmond will find several important uses. Three purposes kept in mind while undertaking this task deserve special mention:

(1) to construct a comprehensive examination in psychology which would be of value to the departmental staff in making recommendations for jobs and graduate schools, (2) to meet the need for an entrance examination for prospective graduate students and special students at the University of Richmond, (3) to provide a standard that would measure the over-all achievement of students majoring in psychology with a view to determining whether they should be candidates for a degree with psychology as the field of concentration.

In order to fulfill these purposes, the writer felt

⁴Mursell, James L. Psychological Testing, p. 223-5.

that the test would have to be of a factorial nature, with a separate score for each test. An omnibus type of test would give only one score, while the Comprehensive Test in Psychology provides for nineteen separate, meaningful scores - certainly an advantage. The test would also have to be long enough so that every subject offered by the department would be properly covered. The test was so designed that it would take the majority of the students about three hours to complete it. The actual median for the initial test group was 160 minutes, with the range 130-202 minutes.

Presented below in Table 1 are the subjects covered by the examination, and the number of items in each sub-test. At the right of each subject is the author (s) of the text (s) used in preparing the questions and answers for that section (See Bibliography A for full information).

Sub-test	N	Text Authors
General	65	Mum; Ruch
Experimental	15	Tinker; Woodworth
Statistics	10	None
History	10	Heidbreder; Murphy
Social	25	Krech & Crutchfield
Child	25	Jersild
Applied	12	Poffenberger
Educational	13	Pressey
Testing	15	Mursell; Wechsler
Projective	15	Bell
Abnormal	14	Conklin
Mental Hygiene	10	Carroll
Personality	10	Murphy; Stagner
Clinical	15	Watson; Wechsler
Industrial	13	Maier
Personnel	12	Bellows

TABLE 1. Subjects and Number of Items of Each Represented on Examination, Plus Authors of Texts Used in Devising Examinations.

The above books were used as guides in forming each sub-test; class notes and the writer's own knowledge provided more material for devising the test items. The answer key was checked by the use of the same texts, by discussion between the writer and staff members, or by discussion with other students.

The initial Comprehensive Test in Psychology was given to thirty-one seniors who were majoring in psychology, on May 3, 1950. The examination was presented as a power test under the Honor System as used at the University of Richmond, i.e., the students were allowed to take breaks and leave the room whenever they chose. The testees were told they could have unlimited time, although after three hours had passed, those few who were still working were encouraged to move with greater speed. Under the above conditions, it was felt that the fatigue and tension of the students would be at a minimum, and that full opportunity would be given them to display their knowledge.

The instructions given the psychology majors were as follows:

You are to try to answer all questions in all sections. Only those answers which are correct will figure in the scoring. There is no time limit. DO NOT WRITE IN TEST BOOKLET. All answers must be indicated on answer sheets. Only one answer will be accepted for each question; select the one best answer which completes the statement or answers the question. No questions may be asked during the examination.

They were also informed that all test items would be of the multiple-choice (four choices) type. This type of question was chosen as it was thought that the multiple-choice statement combined validity-reliability and ease of scoring better than any other kind of test item.⁵

Several days after the examination was administered, forms for estimating their performance on each sub-test were distributed to the testees. On these forms they were asked to indicate, to the nearest decile, the percentage of questions they thought they had right on each subject covered on the test. The analysis of the results follow in Part Three.

Also presented in this paper are the relationships between course grades and test scores (Part One), and the order of difficulty of the subjects represented on the examination (Part Two). In the Conclusion, the validity and reliability of the test will be discussed.

⁵Andrew and Bird, "The Comparative Validity of New Type Questions," J. Educ. Psychology, 1937, 28: 241-258.

Lee, J. M. and Symonds, P. M., "New Type of Objective Tests: A Summary of Recent Investigations", J. Educ. Psychology, 1934, 25: 161-184.

Remmers, H. H., et al., "An Experimental Study of the Relative Difficulty of True-false, Multiple Choice, and Incomplete Sentence Types of Examination Questions", J. Educ. Psychology, 1923, 14: 367-372.

PART ONE

Relationships Between Course Grades and Test Scores

The scores for each sub-test were derived and so-noted on the spaces provided on the first page of each student's answer booklet. Three additional scores were derived, also (See App. A). Each student received a score on the complete test, composed of all sixteen sub-tests; secondly, he was scored on the sum of the sub-tests which were covered by subjects he had taken or was taking at the time; he received a third score on the sum of the sub-tests which were not covered by subjects he had taken. Prior to beginning the test, the testees checked the courses they had taken or were taking on the first page of the answer booklet. Both the score and percentage of items correct for each category (three principle scores) were recorded, although, of course, the latter are more meaningful.

The rank difference method of correlating was used for determining the relationship between course grades and test

scores. Rankings were obtained for each category stated above, with percentages used as a basis, and rankings were obtained for the grades by averaging all grades in psychology (including transferred grades), using the system A-4, B-3, C-2, and D-1.

As may be seen in Table 2, all three correlational coefficients were significantly high, even the lowest being well above the 1% level. Although one might, on "a priori" grounds, expect the highest coefficient to be related to the subjects-taken score, there is a logical explanation for the highest coefficient being obtained for the total score. For it is not unreasonable to find that the better students, as measured by grades in psychology over a four-year period, are more interested in psychology as a whole, and therefore know more than the poorer students about material they have not studied formally. Also, there is probably a greater amount of transfer from one subject to another among the better, more intelligent students. Pressey⁶ indicates this when he states that transfer occurs only when the person is aware of the relatedness of the two subjects. It should be expected that the brighter students would be able to generalize and see similar elements among different subjects better than the poorer students. It is interesting to note

⁶Pressey and Robinson, Psychology and the New Education, p. 571.

that the scores on subjects not taken have a statistically significant relationship with course grades. This shows that a large amount of transfer of knowledge and interest occurs.

All items	Rho =	.85
Subject taken	Rho =	.76
Subjects not taken	Rho =	.66

Table 2. Relationship Between Course Grades and Test Scores.

It was the writer's intention and hope that the test would be able to discern those students who, by cramming, were able to get acceptable grades on tests without retaining much information or understanding of psychology. It was also expected that a few students whose grades, for various reasons, were not of the highest caliber, yet who were thought to understand a great deal about human behavior, would show their worth on the examination. Therefore, much higher coefficients than those obtained were neither wanted nor expected.

As examples of the above mentioned "crammer", the reader may notice that students K, P, and T (see App. A) have a much higher ranking in grades than in test scores. These students, in the writer's estimation, fit perfectly

the first category - crammers.

Students F and G, on the other hand, are examples of those whom the writer feels know more psychology than their grades indicate.

PART TWO

Order of Difficulty of Subjects Represented on Examination.

It would perhaps be desirable for all the sub-tests of a comprehensive examination to be of equal difficulty; yet, such an accomplishment would be almost impossible and at the best would entail a colossal amount of work. So no serious attempt was made to undertake such a task. The order of efficiency for the subjects, therefore, is not too meaningful. (Although the difference in scores between the top and the bottom subjects should be of interest.) This is especially true since five of the sub-tests intentionally contain an "easy" section, these being intended mainly to measure the aptitude of prospective graduate students and special students for specific subjects.

Listed in Table 3 are the sixteen sub-tests comprising the examination. This listing shows two things. First, the number at the left of each subject indicates, in decreasing

order, the success of the testees on the different sub-tests (i.e., the students did best on Child, worst on Personnel). This applies only to those students who had taken the various courses or were taking them. Also in that column is the number of students having taken each course, (i.e., of the thirty-one students taking the examination, twenty-two had taken, or were taking, Child Psychology, and their mean score on the Child sub-test was 68%).

Secondly, Table 3 reveals the same thing for those not having taken the subjects listed. The ranking for the difficulty of the subjects, considering only those students who had not taken the various courses, appears in parentheses at the right of each subject.

Subject	Taken N % correct	Not Taken N % correct
1. Child (1)	68% 22	64% 9
2. Testing (6)	67% 20	48% 11
3. Projective (15)	65% 20	29% 11
4. Clinical (8)	62% 11	41% 20
5. Educational (2)	62% 18	59% 13
6. Social (3)	65% 4	56% 27
7. Abnormal (11)	60% 29	36% 2
8. History (12)	60% 6	34% 25
9. General	58% 31	00 00
10. Industrial (5)	58% 9	54% 22
11. Applied (4)	55% 9	56% 2
12. Experimental (10)	55% 29	37% 2
13. Personality (9)	48% 25	40% 6
14. Statistics	45% 31	00 00
15. Mental Hygiene (7)	45% 14	44% 17
16. Personnel (14)	31% 11	20% 20

Table 3. Order of Success of Subjects Represented on Comprehensive Examination.

Possibly of greater interest than the relative success or difficulty rank of the subjects is the comparison of the average score of those who have taken a subject with the score of those who have not formally studied the subject. In Table 3, it can be seen that there are five subjects whose average scores show no appreciable difference. At least four of these could be expected, so such results are not surprising. Perhaps the test is faulty in not discriminating between those who had taken the subjects and those who had not; but, then, one might expect psychology seniors to possess a reasonable knowledge about such subjects as Child and Educational Psychology, Applied Psychology, and Mental Hygiene, even without formal work in the courses. There is much overlap between courses in Child and Educational Psychology, and a student who has had just one of the two should do relatively well on a test on the other subject, especially if he has studied other related subjects. The material presented in a Mental Hygiene Course is also partially discussed in other courses such as Abnormal, Personality, and Clinical. Applied is covered partially in other courses (though this explanation is not satisfactory, since for this section the writer attempted to choose questions which would not involve much transfer from other courses) but the Industrial scores cannot be similarly explained.

In a further endeavor to try to explain the closeness of the scores for those who had taken and for those who had not taken the above courses, the writer decided to use a statistical approach. Separating the two groups of students for each of the five courses, the writer calculated the mean rank (using ranks for the complete test) for the students who had taken and for the students not having taken each course. The question was, were the two groups for each course of equal ability? If those who had taken the course were of less general ability than those who had not, then that would be at least a partial explanation, itself, as to the closeness of the groups' scores. The results may be seen in Table 4.

Course	<u>Mean Rank</u>		<u>Mean Rank</u>	
	Taken	N	Not Taken	N
Mental Hygiene	20.0	14	12.8	17
Educational	17.4	18	14.1	13
Child	16.1	22	15.7	9
Applied	15.3	9	16.3	22
Industrial	15.8	9	15.1	22

Table 4. Comparison of Groups Which Took Subject With Group Which Did Not - In Regard to Rank on Complete Test.

It is well evident that those who had not taken Mental Hygiene (N=17) were far superior on the whole examination,

and presumably better students, than those who had taken the course (N=14). Thus, it should be understandable that those who never formally took a course in Mental Hygiene still knew just as much about the subject. The same reasoning (though not having as strong a basis) may apply to the Educational sub-test, for those not having taken the course seem to be better students, as the mean rank for those students is 14.1, while those who did take Educational Psychology have a mean rank of 17.4.

The close scores on the other three courses, however, are not as easily explained. Child shows a slight tendency toward indicating a positive difference in ability between the two groups, but Industrial and Applied show a negative trend. The latter two sub-tests may need revising, but the small number of students having taken each course does not insure reliable interpretation.

The number of students having taken or not having taken other courses are not equally distributed; therefore, rather than showing the gain in knowledge obtained by taking a course, Table 4 may be indicating, to some degree, the difference in over-all ability of the different groups of students.

Finally, three more of the scores are not strictly comparable (Social, Experimental and Abnormal), since

comparisons are meaningless with only two or four cases in one or the other categories.

It might be noted here that the Statistics sub-test consisted only of elementary statistics encountered in the entire field of psychology as offered at the University of Richmond, since no formal course is offered by the Psychology Department at this time (though some of the students have taken a course in Statistics from the Department of Mathematics). The low score on that test (see Table 3) is enlightening, in that it indicates the need for more stress on statistics in various psychology courses. More preferable, of course, would be a separate course in psychological statistics, of which there is a possibility, if conditions permit.

Not particularly pertaining to an explanation of the examination itself, but of possible interest to the reader, is Table 5. In this table are listed the psychology subjects offered at the University of Richmond, with "N" representing the number of students - considering only the thirty-one seniors who took the Comprehensive Test in Psychology - who received credit for each course.

Subject	N
* General	31
* Experimental	29
Abnormal	29
Personality	26
Child	22
Projective	20
Testing	19
Educational	18
Mental Hygiene	14
Personnel	11
Clinical	11
Applied	9
Industrial	9
History	6
Social	4
*Required	

Table 5. Subjects in Order of Requirement or Election.

PART THREE

Students' Estimations of Their Test Performances

As was stated in the introduction, the students indicated on a form their estimation of the percentage of items correct on each sub-test. In computing the accuracy of these estimations, the difference between the estimation and the actual score was noted, using the nearest decile point as a base. Although this measurement is somewhat rough, it was assumed that errors would cancel out, since there was a total of 432 computations.

Analysis of the twenty-seven sets of estimations returned showed the following results:

The mean underestimation was .86 deciles, or 8.6%.

The mean overestimation was .94 deciles, or 9.4%.

The above resulted in a net mean underestimation of .06 deciles.

The mean deviation, considering the sum of under and overestimation for each person, was 1.61 deciles, or 16.1%.

Comparing the upper and lower quarters (using total scores), the writer found that the lower quarter overestimation (.63 deciles) was about the same degree as the upper quarter underestimation (.60 deciles), and that the lower quarter had a statistically significant greater deviation (1.90 deciles) than the upper quarter (1.57 deciles).

The data collected on estimation of ability agrees with earlier studies,⁷ which also found that brighter students tended to underestimate their ability, while the poorer students tended to overestimate. Of further interest was the finding that the poorer students are not able to judge their performances as well as the better students. Whether underestimating or overestimating, they deviate more widely from the actual result than the brighter students. Sears⁸ also found higher discrepancy between aspired goal and

⁷Anderson, H. H. and Brandt, H. F., "Study of Motivation Involving Self-announced Goals of 5th Grade Children and the Concept of Level of Aspiration," J. Social Psychology, 1939, 10: 209-252.

Hilgard, E. R., Sait, E. M., and Margaret, G. A., "Level of Aspiration as Affected by Relative Standing in an Experimental Social Group". J. Experimental Psychology, 1940, 127: 411-421.

⁸Sears, P. S., "Levels of Aspiration in Academically Successful and Unsuccessful Children, J. Abnormal and Social Psychology, 1940, 35: 498-536.

reality for the group with a record of failure than for the
successful, highly confident group.

CONCLUSION

Although the sample of testees is presently not large enough for fully determining the validity of the Comprehensive Examination, the writer's opinion, at the moment, is that the purposes of the examination seem to have been fulfilled.

The test seems to be of sufficient difficulty to discriminate among the good, average, and poor students. Objective evidence for this is offered by the mean scores and standard deviations given below in Table 6.

	<u>M</u>	<u>S.D.</u>
Subjects Taken	57%	10.2
Subjects Not Taken	47%	9.6
Total	54%	8.5

Table 6. Mean Scores and Standard Deviations.

The above data shows that the scores approximate normal curves as closely as one could desire, considering the small group involved. Ross⁹ stated that the average student should get about half of the items of an objective test correct, the closer to fifty per cent, the better. The Comprehensive Test in Psychology meets that requirement. Ross also offers a number of other suggestions¹⁰ for the construction of a good objective test which the writer has followed. These are: (1) "Avoid making the correct response consistently longer or shorter than the others;" (2) "Make all responses plausible;" (3) "At least four choices should be presented whenever possible;" (4) "In phrasing multiple-choice test items, consideration should be given to the fact that the answer may be arrived at by eliminating the incorrect responses as well as by selecting the correct response directly. Requiring the pupil to select the least satisfactory response in the series given, or one that is not true, will often compel a careful comparison of all the possible responses;" (5) "Require the simplest possible method of indicating the response;" (6) Use the correction formula only if the number of choices is fewer than four;" (7) "Make

⁹Ross, C. C., Measurement in Today's Schools, p. 124.

¹⁰Ibid, p. 124-5.

all optional responses grammatically consistent;" (8)"Avoid using in the correct response the same words or phrases that occur in the question or incomplete statement."

In discussing the evaluation of a test, Ross offers five principles which he believes should be followed. This writer used three¹¹ of these guides; discussion of the fourth, the desirability of obtaining a reliability coefficient, follows.

Because of the difficulties involved, no attempt was undertaken to determine the reliability of the test. At the present time, there is only one form of the examination, which leaves two possible methods of finding the reliability. One, the test - retest method, involves several uncontrollable variables such as discussion among the students between testing periods, and extra studying. Also, since the test is relatively long, some time would have to pass between trials. This would create an awkward situation, as the students should take the test in May in order that they have sufficient knowledge of the courses being taken, while during the same month the regular final examinations occur. The attitudes of the students toward taking two three-hour tests within a week, then enduring two weeks more of examinations, can be imagined!

¹¹(1) "The difficulty of a test is a rough indication of its validity;" (2) Items should be criticized by persons taking the test; (3) Results of the test should be checked against an outside criterion.

The other method of determining reliability, split-half (or even-odd) would be impossible to use with this factorial type of test. Not only do the items for each sub-test vary in difficulty, but the different sub-tests vary among themselves.

So, rather than try some hit-or-miss means of securing a reliability coefficient, the writer decided that, for the moment at least, no quantified reliability could possibly be obtained. There is reason to believe, however, that for such a test as this, the reliability should be sufficiently high. For the test questions appear (to the writer and the testees) to be easily understood, and ambiguity seems to have been held to a minimum.

The students' scores, with two or three exceptions, agreed closely with subjective estimates retained by the writer and the psychology staff. This is further evidence of the test's initial validity.

Further checks on the test's validity, such as the comparison of graduate school grades and the scores on the psychology section of the Graduate Record Examination with the Comprehensive Test in Psychology scores, will have to await the passage of time.

Generally speaking, the writer feels that his designated objectives were accomplished. Not only has the test

been of use to the Psychology Department, but the students who took the Comprehensive Test in Psychology were able to determine their strengths and weaknesses in psychology, and also to find how they compared with their peers. It is hoped that further use of the examination will prove valuable to all concerned.

APPENDIX A
Chart of Students' Performances

Students	Subjs Taken		Subjs. not taken		Total		Ranks			Grade avg.	Rank
	raw	%	raw	%	raw	%	taken	not	tot.		
A	187	<u>78</u>	24	<u>63</u>	211	<u>76</u>	1	1	1	3.8	1
B	146	<u>75</u>	51	<u>60</u>	197	<u>71</u>	2	2.5	2	3.6	2
C	151	<u>73</u>	43	<u>60</u>	194	<u>70</u>	3	2.5	3	3.2	5
D	136	<u>70</u>	49	<u>59</u>	185	<u>66</u>	4	4.5	4	2.50	13.5
E	111	<u>65</u>	63	<u>59</u>	174	<u>62</u>	6	4.5	5	3.5	3
F	108	<u>63</u>	61	<u>57</u>	169	<u>61</u>	7	6.5	6.5	2.44	15
G	129	<u>66</u>	41	<u>49</u>	170	<u>61</u>	5	10.5	6.5	2.80	9
H	120	<u>62</u>	41	<u>48</u>	161	<u>58</u>	9	12.5	8	2.7	10
I	121	<u>62</u>	39	<u>47</u>	160	<u>57</u>	9	16	9	3.1	6
J	67	<u>56</u>	88	<u>55</u>	155	<u>56</u>	16	8	11	2.50	13.5
K	124	<u>59</u>	33	<u>47</u>	157	<u>56</u>	13	16	11	2.9	7
L	107	<u>62</u>	50	<u>46</u>	157	<u>56</u>	9	19	11	2.55	11.5
M	125	<u>56</u>	27	<u>47</u>	152	<u>54</u>	16	16	13	2.83	8
N	96	<u>53</u>	48	<u>49</u>	144	<u>53</u>	21.5	10.5	14.5	2.33	19.5
O	107	<u>59</u>	40	<u>41</u>	147	<u>53</u>	13	23	14.5	2.33	19.5
P	93	<u>55</u>	53	<u>49</u>	146	<u>52</u>	19	12.5	17.5	3.3	4
Q	80	<u>49</u>	66	<u>57</u>	146	<u>52</u>	27	6.5	17.5	2.38	17
R	98	<u>60</u>	47	<u>40</u>	145	<u>52</u>	11	24.5	17.5	2.38	17
S	87	<u>51</u>	59	<u>54</u>	146	<u>52</u>	25.5	9	17.5	2.17	23.5
T	116	<u>56</u>	26	<u>36</u>	142	51	16	28	21	2.55	11.5
U	93	<u>55</u>	50	<u>45</u>	143	<u>51</u>	19	21	21	2.38	17

Students	Subjs Taken		Subjs. not taken		Total		Ranks			Grade avg.	Rank
	raw	%	raw	%	raw	%	taken	not	tot.		
V	92	<u>55</u>	51	<u>46</u>	143	<u>51</u>	19	19	21	2.2	23
W	79	<u>51</u>	58	<u>46</u>	137	<u>49</u>	25.5	19	24	2.12	25
X	102	<u>59</u>	35	<u>33</u>	137	<u>49</u>	13	31	24	1.78	31
Y	100	<u>52</u>	38	<u>44</u>	138	<u>49</u>	23.5	22	24	2.30	21
Z	106	<u>53</u>	29	<u>37</u>	135	<u>48</u>	21.5	26.5	26	1.82	30
AA	81	<u>46</u>	48	<u>47</u>	129	<u>46</u>	29	16	27	1.89	28
BB	84	<u>48</u>	37	<u>35</u>	121	<u>43</u>	28	29.5	29	2.0	26.5
CC	76	<u>52</u>	50	<u>37</u>	126	<u>45</u>	23.5	26.5	29	2.17	23.5
DD	48	<u>33</u>	53	<u>40</u>	101	<u>36</u>	31	24.5	30.5	1.88	29
EE	67	<u>36</u>	33	<u>35</u>	100	<u>36</u>	30	29.5	30.5	2.0	26.5
↓	31										
		↓	↓	↓	↓					↓	
		M = 57%	M = 47%		M = 54%					2.53	
		S.D. = 10.2	S.D. = 9.6		S.D. = 8.5						

APPENDIX B
Summary of Calculations

A. Calculations Used in Finding Standard Deviations of Test Scores

$$\sigma = \sqrt{\frac{\sum X^2}{N} - M^2}$$

$$\sigma = \sqrt{\frac{103980}{31} - (57)^2}$$

$$\sigma = 10.2$$

$$\sigma = \sqrt{\frac{71345}{31} - (47)^2}$$

$$\sigma = 9.6$$

$$\sigma = \sqrt{\frac{92622}{31} - (54)^2}$$

$$\sigma = 8.5$$

B. Calculations Used in Finding Relationships Between Course Grades and Test Scores

$$R_{HO} = 1 - \frac{6SD^2}{N(N^2-1)}$$

$$R_{HO} = 1 - \frac{6(745.00)}{31(960)}$$

$$R_{HO} = .85$$

$$R_{HO} = 1 - \frac{6(1187.50)}{31(960)}$$

$$R_{HO} = .76$$

$$R_{HO} = 1 - \frac{6(1650.50)}{31(960)}$$

$$R_{HO} = .66$$

C. Calculations Used in Determining Significance of Students' Estimations of Their Test Performances

$$\sigma = \sqrt{\frac{\sum X^2}{N} - M^2}$$

$$\sigma_M = \frac{\sigma}{\sqrt{N}}$$

$$\sigma_x = \sqrt{\frac{20.34}{7} - (1.57)^2}$$

$$\sigma_{M_x} = \frac{.21}{2.64}$$

$$\sigma_x = .21$$

$$\sigma_{M_x} = .080$$

$$\sigma_y = \sqrt{\frac{27.45}{7} - (1.90)^2}$$

$$\sigma_{M_y} = \frac{.18}{2.64}$$

$$\sigma_y = .18$$

$$\sigma_{M_y} = .068$$

$$\pi = \frac{M_x - M_y}{\sqrt{\sigma_{M_x}^2 + \sigma_{M_y}^2}}$$

$$\pi = \frac{1.57 - 1.90}{\sqrt{(.080)^2 + (.068)^2}}$$

$$\pi = 3.3$$

BIBLIOGRAPHY A

Texts Used in Forming Test Items

Bell, John E., Projective Techniques, N.Y.: Longmans, Green and Co., 1948.

Bellows, Roger H., Psychology of Personnel in Business and Industry, N. Y.: Prentice-Hall, 1949.

Carroll, Herbert A., Mental Hygiene, N.Y.: Prentice-Hall, 1947.

Conklin, Edmund S., Principles of Abnormal Psychology (revised), Henry Holt and Co., 1936.

Heidbroder, Edna, Seven Psychologies, N.Y.: Appleton-Century-Crofts, Inc., 1947.

Jersild, Arthur T., Child Psychology, N.Y.: Prentice-Hall 1947.

Krech, David and Crutchfield, Richard S., Theory and Problems of Social Psychology, N.Y.: McGraw-Hill Book Co., Inc., 1947.

Krier, Norman, R.F., Psychology in Industry, Boston: Houghton Mifflin Co., 1946.

Munn, Norman L., Psychology, Boston: Houghton Mifflin Co., 1946.

Murphy, Gardner, Historical Introduction to Modern Psychology, N. Y.: Harcourt, Brace and Co., 1949.

- Murphy, Gardner, Personality, N. Y.: Harper and Brothers, 1947.
- Mursell, James L., Psychological Testing, N. Y.: Longmans, Green and Co., 1949.
- Poffenberger, A. T., Principles of Applied Psychology, N.Y.: D. Appleton-Century Co., Inc., 1942.
- Pressey, Sidney L. and Robinson, Francis P., Psychology and the New Education (revised), N. Y.: Harper and Brothers, 1944.
- Ruch, Floyd L., Psychology and Life, Chicago: Scott, Foresman and Co., 1941.
- Stagner, Ross, Psychology of Personality, N.Y.: McGraw-Hill, 1937.
- Tinker, Miles A., Introduction to Methods in Experimental Psychology, N. Y.: D. Appleton-Century Co., 1947.
- Watson, Robert I. (editor), Readings in the Clinical Method in Psychology, N. Y.: Harper and Brothers, 1949.
- Wechsler, David, The Measurement of Adult Intelligence, Baltimore: The Williams and Wilkins Co., 1944.
- Woodworth, Robert S., Experimental Psychology, N. Y.: McGraw-Hill, 1937.

BIBLIOGRAPHY B

References Used in Writing This Paper

- Anderson, H. H., and Brandt, H. P., "Study of Motivation Involving Self-Announced Goals of Fifth Grade Children and the Concept of Level of Aspiration", J. Social Psychology, 1939, 10: 209-232.
- Buros, Oscar Krisen, Editor, Mental Measurements Yearbook (1940), Arlington: The Gryphon Press, 1945.
- Falls, J. D., "Research in Secondary Education", Kentucky School Journal, 1928, 6: 42-46.
- Gould, R., "Some Sociological Determinants of Goal Strivings", J. Social Psychology, 1941, 13: 461-473.
- Hilgard, E. R., Sait, E. M., and Margaret, G. H., "Level of Aspiration as Affected by Relative Standing in an Experimental Social Group", J. Experimental Psychology, 1940, 127: 411-421.
- Hulten, G. E., "The Personal Element in Teachers' Marks", J. of Educational Research, 1925, 12: 49-55.
- Hunt, J. McV., Personality and the Behavior Disorders, vol. 1, Chap. 10, N. Y.: The Ronald Press Co., 1944.
- Pressey, Sidney L. and Robinson, Francis P., Psychology and the New Education (revised), N. Y.: Harper and Brothers, 1944.

- Ross, C. C., Measurement in Today's Schools, N. Y.: Prentice-Hall, Inc., 1947.
- Sears, P. S., "Levels of Aspiration Academically Successful and Unsuccessful Children," J. Abnormal and Social Psychology, 1940, 35: 498-536.
- Starch, Daniel and Elliot, E. C., "Reliability of Grading Work in Mathematics", School Review, 1913, 21: 254-259.

VITA

The writer was born in Norfolk, Virginia, November 1, 1926. Educated in the public schools of Norfolk and Newport News, he was graduated from the Newport News High School in 1943.

After serving in the United States Air Corps, 1945-6, he entered Mary Washington College of the University of Virginia for his freshman year. At the University of Richmond, in September, 1949, he received the Bachelor of Arts degree in Psychology; and then he entered the Graduate School of the University.

The writer plans to continue his graduate work at Vanderbilt University.