

1955

# Predictive measurements of college success : a study of the relationship between the freshman testing program and college success at the University of Richmond

Walter Nelson Taylor

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



Part of the [Psychiatry and Psychology Commons](#)

---

## Recommended Citation

Taylor, Walter Nelson, "Predictive measurements of college success : a study of the relationship between the freshman testing program and college success at the University of Richmond" (1955). *Master's Theses*. 1334.  
<https://scholarship.richmond.edu/masters-theses/1334>

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](mailto:scholarshiprepository@richmond.edu).

PREDICTIVE MEASUREMENTS OF COLLEGE SUCCESS

A STUDY OF THE RELATIONSHIP BETWEEN  
THE FRESHMAN TESTING PROGRAM AND COLLEGE SUCCESS  
AT THE UNIVERSITY OF RICHMOND

BY

WALTER NELSON TAYLOR

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

AUGUST, 1955

LIBRARY  
UNIVERSITY OF RICHMOND  
VIRGINIA

## TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION . . . . .	1
Purpose of this Study . . . . .	2
Need for such a Study . . . . .	3
II. A REVIEW OF PREVIOUS RESEARCH . . . . .	5
Literature on the A.C.E. . . . .	6
Reliability . . . . .	6
Validity . . . . .	8
In terms of college marks . . . . .	8
Compared with other tests . . . . .	11
A.C.E. scores and other predictive measures.	13
The A.C.E. as a differential predictor . . . .	15
Factors which may affect A.C.E. scores and	
validity . . . . .	18
Admissions policy . . . . .	18
Age and previous military status . . . . .	19
Sex . . . . .	20
Other factors . . . . .	20
Literature on the Cooperative Reading Test . . .	21
Reliability . . . . .	22
Validity . . . . .	23

CHAPTER	iv PAGE
III. METHODS AND PROCEDURE . . . . .	25
Choice of Criterion . . . . .	25
Subjects . . . . .	27
The tests employed . . . . .	30
Procedure . . . . .	31
General . . . . .	31
Statistical method employed in obtaining correlations . . . . .	32
IV. RESULTS . . . . .	34
Summary of Results . . . . .	41
Results Compared with Those Published . . . . .	44
V. SUMMARY AND CONCLUSIONS . . . . .	48
Summary . . . . .	48
Conclusions . . . . .	51
BIBLIOGRAPHY . . . . .	57
APPENDIX . . . . .	62
VITA . . . . .	67

## LIST OF TABLES

TABLE		PAGE
I.	Students Eliminated from Subject Sample By Reasons which Disqualified Them . . . . .	28
II.	Number of Students Eliminated Under Automatic Ruling, By Semesters, Class of 1954 . . . . .	29
III.	Consolidated Summary of Data . . . . .	35
IV.	Correlations of A.C.E. Scores Q, L, T and Cooperative Reading Test Total Scores with College Success at University of Richmond . . .	36
V.	Comparison of Mean Scores for Nation, Class of 1954, and Study Group . . . . .	36
VI.	Frequency Distribution Table for Total A.C.E. Scores . . . . .	62
VII.	Frequency Distribution Table for L Scores . . .	63
VIII.	Frequency Distribution Table for Q Scores . . .	64
IX.	Frequency Distribution Table for Total Reading Scores . . . . .	65
X.	Sample Calculation of Biserial Coefficient of Correlation, Based on Data from Table VI . . .	66

## LIST OF FIGURES

FIGURE	PAGE
1. Pass Group: Distribution of Scores on Cooperative Reading Test . . . . .	37
2. Fail Group: Distribution of Scores on Cooperative Reading Test . . . . .	37
3. Pass Group; Distribution of Scores on A.C.E. Total.	38
4. Fail Group; Distribution of Scores on A.C.E. Total.	38
5. Graph of Scores Obtained on Cooperative Reading Test by Pass and Fail Groups . . . . .	39
6. Percentage Failure Rate of Students Scoring at or Below a Given Percentile on The Cooperative Reading Test. Percentiles are those for University of Richmond, 1949 . . . . .	40

## CHAPTER I

### INTRODUCTION

From its birth with the Binet Intelligence Scale, half a century ago, mental testing has assumed a place of continually growing importance in the worlds of education and psychology. As is true of any instrument, the more one learns of its uses and possibilities, the more one begins to recognize its limitations and inadequacies. Thus, while mental testing has become firmly established as a procedure of undeniable consequence and utility in the prediction of academic success, guidance, and the like, experience and research have made it plain that its usefulness depends upon strict observation of the instrument's limitations. All too often the test user is prone to put more emphasis upon test scores than is warranted in light of the true reliability and validity. Carried to its logical extreme, this usually results in an inappropriate and premature condemnation of the test in question, as well as a loss of faith in tests in general. All too often we fail to "test the test." Unknown limitations cannot be observed,

## 1. PURPOSE OF THIS STUDY

The writer undertook this study with several objectives in view. Primarily it was desired to study the exact statistical relationship between academic success at University of Richmond and each of several scores obtained in the freshman testing program. More specifically, it was desired to check the predictive ability of: (1) the "Q" or quantitative score, (2) the "L" or linguistic score, (3) the "T" or total score, (sum of "Q" and "L" scores) obtained from the American Council on Education Psychological Examination for College Freshmen. Hereinafter, this test will be referred to as the A.C.E.

In addition to this, it was desired to compute the validity coefficient for another test of the battery, the Cooperative Reading Comprehension Test. This is one of three complete tests which are combined in the testing program to make up the Cooperative English Test. Hereinafter, this sub-test score will be referred to as the Total Reading score.

Having obtained the above information, which could be considered an end in itself, the writer then would be in the position to investigate the hypothesis that reading ability might well prove to be as important to college success as is the quality which is measured on the A.C.E., usually thought of as native mental ability.

Finally, it was the purpose of this study to review



the literature for reports of previous investigations of this and closely related problems. This was done not only as a check on the writer's findings, but to prevent him from covering over-investigated territory. Also, it was a search for an indication of the most profitable direction of exploration.

A partial review of the pertinent literature will be presented in the next chapter.

## II. NEED FOR SUCH A STUDY

Speaking of tests in general, Mursell admonishes:

It is always dangerous to assume that a mental test can reveal or measure intelligence, aptitude, or talent ...or can uncover its universal essence. It can only reveal and deal with any such function or trait in the setting of a particular population...<sup>1</sup>

This appears to be the key to the problem cited at the beginning of this chapter for tests of all kinds. It is particularly true for the A.C.E. and the Cooperative Reading Test. Many of the investigators reviewed noted the wide variety of results in different institutions and recommended that each college obtain its own coefficient of correlation. Some of those making the most extensive investigations were

---

1. James L. Mursell, Psychological Testing (New York: Longmans, Green and Company, 1947), p. 63

Wallace<sup>2</sup>, MacPhail<sup>3</sup>, Smith and Triggs<sup>4</sup>. These investigations will be reported in the next chapter.

In addition to "testing the test", it was hoped that such a study might provide data which could be worked into a critical score for selection of college entrants, or entrants of specific courses. Certainly this would provide an improved basis for future counseling.

Finally, this investigation should provide a basis for further study, and, if nothing else, at least point the way. One thing appears certain in the light of nationwide findings. We are not justified in relying upon either of these tests until we have some measure of their efficacy as they apply to a particular population sample.

---

2. W. L. Wallace, "Differential Predictive Value of The A.C.E. Psychological Examination," School and Society, 70:23-5, July 9, 1949.

3. Andrew H. MacPhail, "Q and L Scores on the A.C.E. Psychological Examination," School and Society, 56:248-51, Sept. 19, 1942.

4. D. D. Smith and Frances O. Triggs, "Educational Successes and Failures of Students with High "Q" and Low "L" Scores on the A.C.E. Examination," American Psychologist, 5:353-4, July, 1950.

## CHAPTER II

### A REVIEW OF PREVIOUS RESEARCH

A great deal has been written in regard to the A.C.E. and the Cooperative Reading Test. The literature on the A.C.E. is, in fact, quite extensive. However, it is beyond our scope and purpose to present all of it here. Here the writer will attempt to give only a brief summary of some of the most extensive and significant studies done on problems very closely related to the one at hand. Also, an attempt has been made to confine ourselves to those studies of comparative recency. For the reader's convenience, references to some of the works not included here will be cited in the Bibliography.

When interpreting the results of research on these tests, the reader should bear in mind that the Freshman Testing Program was participated in by 144 colleges, and 20,470 students located in 40 states<sup>1</sup>. Naturally we should expect a great deal of fluctuation in the findings. Extreme variations attributable to type of institution, geographical region, type of control, coed and male and female attendance, number of students tested

---

1. Educational Testing Service, Final Report on the 1952 National College Freshman Testing Program, (Cooperative Test Division, Educational Testing Service Report), P.1.

in a particular study and other variables should each be expected to have its effect upon the results. Such was the case in this instance. The Educational Testing Service reports an even greater diversity of results than this writer would have expected<sup>2</sup>. It is significant that the test publishers provide no national validity or reliability studies for this reason. Such figures would have no meaning for a particular institution. They do, however, provide a set of national norms based upon all of the participants each year which is a helpful yardstick for checking the standing of one's own institution against a very similar one.

#### I. LITERATURE ON THE A.C.E.

Since the research on the A.C.E. is quite extensive it will be necessary to divide the reports into sub-groups according to the phase with which they are primarily dealing.

##### Reliability

The reliability of the A.C.E. appears to be quite respectable. Angoff employed a modified test-retest approach to reliability using the 1949 edition (the same test employed in our research) and found reliability coefficients of .89 to .92 for Q, .90 to .94 for L, and .93 to .96 for Total<sup>3</sup>. Since

---

2. Ibid., passim.

3. W. H. Angoff, "Test Reliability and Effective Test Length," Psychometrika, 18:1-14, 1953.

Angoff used 666 and 787 students in his study we can be reasonably certain that his results were not obtained through a spurious sample. Super reports that the reliability has been consistently high regardless of which form is tested.<sup>4</sup> He reports an odd-even reliability of .95. Mursell states that one of the great values of the A.C.E. is that it is revised yearly, and that since 1940 an analysis of item difficulty has been set up so that scores of successive editions are comparable.<sup>5</sup>

Samenfeld conducted a study wherein he gave the A.C.E. to a group of high school freshmen, and retested the same group when they were high school seniors.<sup>6</sup> Checking their two sets of scores against college grades, he found that the first correlated as well as the second set made when they were graduating seniors. This not only attests the reliability of the test and its successive forms, but it points out an interesting sidelight. From these findings it would appear that high school freshmen might be tested with as much accuracy as seniors, with the added advantage of having counseling information four years in advance.

---

4. Donald E. Super, Appraising Vocational Fitness By Means of Psychological Tests, (New York: Harper and Brothers, 1949), p. 117.

5. Mursell, op. cit., p. 161.

6. Herbert W. Samenfeld, "Predicting College Achievement," Journal of Higher Education, 24:432-3, November, 1953.

## Validity

In Terms of College Marks. This is the area in which the most extensive research has centered. As was to be expected, the results are varied and the problem has been approached from many standpoints.

Cronbach reports an R as high as .57 for A.C.E. Total Vs. college freshman marks.<sup>7</sup> His study employed 97 subjects. Remmers, Elliott and Gage correlated the A.C.E. with first semester grades for 1,981 freshmen and obtained correlations of .36, .38, and .41 for Q, L, and T scores respectively.<sup>8</sup>

Fredericksen and Schrader studied 16 colleges to get a broad and varied sample.<sup>9</sup> Using freshman grades as criteria they obtained a mean coefficient of .47 for total A.C.E.

In a review in Buros' Yearbook of Mental Measurement, Guilford says that the test employs a wide range of items and taps a wider range of abilities than most of its kind, but that it still does not measure too well the things by which achievement is assessed in college.<sup>10</sup>

7. Lee J. Cronbach, Essentials of Psychological Testing,  
(New York: Harper and Brothers, 1949), p. 58.

8. H. H. Remmers, D. N. Elliott, N. L. Gage, "Curricular Differences in Predicting Scholastic Achievement; Applications to Counseling," Journal of Educational Psychology, 40:385-94, Nov. '49

9. Norman Fredericksen and Schrader, W. B., "The A.C.E. Psychological Examination and High School Standing as Predictors of College Success," Journal of Applied Psychology, 36:261-5, Aug., 1954.

10. Oscar K. Buros, (ed.) Third Mental Measurements Yearbook,  
(New Brunswick, N. J.: Rutgers University Press, 1949), p.217.

Wallace compared Q and L scores with freshman, junior and sophomore grades.<sup>11</sup> The highest correlation he could obtain was .49 when a combination of Q and L were run against English grades. He urged that colleges be conservative in applying test results for admission and prediction.

The Educational Testing Service, publishers of the test, report a variety of findings on A.C.E. validity.<sup>12</sup> They show that Berdie, at the University of Minnesota, studied nine colleges and obtained correlation coefficients between A.C.E. scores and college grades ranging from .39 at University of Chicago to .62 at State University of Iowa. Of the nine studies reported, six showed R's between .50 and .60. The Educational Testing Service cites another study in which first term grades for freshmen at eight colleges Vs. A.C.E. scores yielded correlations of from .28 to .49 with a median correlation of .42.<sup>13</sup> This study was conducted by the Testing Service. They report also a study conducted by the Testing Service in which Frederiksen, et. al., obtained a median correlation of .42 for

---

11. Wallace, op. cit., p. 24.

12. Educational Testing Service, "Summary of Selected Research on The Validity of The American Council Psychological Examination For College Freshmen As A Predictor of College Grades," Research Memorandum, (1954), p. 2.

13. Ibid., p. 3.

five groups of students at the University of Rochester.<sup>14</sup> The Educational Testing Service concludes that, in general, A.C.E. total scores correlate with freshman grades in the .40's<sup>15</sup>; whereas, Super, summarizing some additional studies, finds that A.C.E. scores and four-year grades generally correlate around .45<sup>16</sup>.

Generally, then, one would be on fairly safe ground to conclude that A.C.E. scores tend to correlate with college marks somewhere in the mid-forties on the average, or, to be more specific, one would expect a median correlation in this range.

On the other hand, it is also apparent from the studies reviewed that the range of correlations for colleges is very broad, running from the low .30's to the mid .60's.

One fact which is significant as far as this study is concerned is that, in all the literature reviewed by this writer, there is apparently no study recorded in which the criterion was graduation or failure in college. The significance of this will be discussed in a later chapter.

---

14. Ibid., p. 4.

15. Ibid., p. 5.

16. Super, op. cit., p. 120.



Compared with Other Tests. There are few psychometrists so bold as to state that they know just what a test measures. The best one can state is that a given test is constructed in a way calculated to measure a given trait or quality, and that its scores show a definite mathematical relation to some other criterion which is assumed to be an index of the desired trait. When experience and study lead us to believe that a test appears to measure some quality particularly well and consistently, (e.g., the Wechsler-Bellevue Intelligence Scale measuring that nebulous quality called "intelligence",) we often like to compare it with another test to ascertain the latter's tendency to measure this same thing. It is obvious, of course, that one has still not proven just what the new test measures, but only how well it measures what the original test measures.

A fair amount of research has been devoted to correlating the A.C.E. with other measures, as well as to comparing the correlation coefficients obtained by running scores from two tests against the grades of one group of subjects.

Traxler, in two correlations of scores on the Otis Self-Administering Test of Mental Ability and the A.C.E. obtained correlations of .78 and .82.<sup>17</sup> Although Traxler calls the Otis a test of academic aptitude, this test is commonly

---

17. A. E. Traxler, "Correlations Between Two Tests of Academic Aptitude," School and Society, 61:383-4, 1945.

referred to as one of intelligence. Generally speaking, the Otis is thought of as a fairly well standardized test of intelligence, a little easy on the adult level, which correlates with college grades in the .50's and with the Binet at about the same level.<sup>18</sup> Are we measuring "intelligence" or ability in matters academic?

Edward Anderson, and others, correlated the A.C.E. and the Wechsler-Bellevue with college grades made in the freshman year by women.<sup>19</sup> The median correlation between the A.C.E. and grades was .52, while the Wechsler Verbal correlated .52 with grades, the Performance .23 and the Full Scale .45 with college grades. The same group also correlated the A.C.E. scores directly with the Wechsler and found that the 1941 form correlated the higher, with coefficients of .52 Full Scale, .31 Performance and .54 Verbal.

Remmers, in a study already cited, found the Perdue English test a slightly better predictor of freshman grades than the A.C.E.<sup>20</sup> (.41 compared to .47)

---

18. Super, op. cit., pp. 107-114.

19. Edward E. Anderson, et. al., "Wilson College Studies in Psychology: I. A Comparison of the Wechsler-Bellevue, Revised Stanford-Binet, and American Council on Education Tests At The College Level," Journal of Psychology, 14:317-26, 1942.

20. Remmers, Elliot and Gage, op.cit., p. 393.

A.C.E. Scores and Other Predictive Measures. It is well known that there are other measures of performance which are often just as predictive of college success as are formal tests such as the A.C.E. In fact, some are even more predictive at times. One of the most outstanding of these predictive indices is high school grades or class standing. Of course, in order to be useful, the high school grades must be converted to a standard scale taking into account the size of the class, and also other factors must be equated.

Frederiksen has done considerable work in this area. In a study cited above he correlated high school grades and A.C.E. scores with college freshman grades.<sup>21</sup> The A.C.E. correlated .47, high school grades correlated .57 and multiple composites of both indices correlated .68. This study included 16 schools.

Samenfeld, in a similar experiment, chose 186 college students at random, and looked up their high school percentile ranks and A.C.E. scores.<sup>22</sup> High school rank correlated .58 with college grades, while A.C.E. scores correlated only .39 at the highest. A combination of the two, however, yielded a correlation coefficient of .63. Cronbach found that combining high school grades with A.C.E. scores raised the correlation

---

21. Frederiksen and Schrader, op. cit.

22. Samenfeld, op. cit.

Coefficient obtained from A.C.E. scores alone .57 to .79.<sup>23</sup>

Thus it would appear that one could safely expect the multiple correlation coefficient combining high school grades and A.C.E. scores to be in the vicinity of .60 plus.

The above findings appear to substantiate to some degree the writer's assumption that the A.C.E. does measure a good deal of that which we term intelligence when employed within a limited population such as high school and college students. This reasoning is postulated upon the assumption that high school grades, when taken as a predictive index, constitute nothing more or less than a prolonged work sample. They certainly indicate a good deal more than merely an individual's position in relation to his peers. Indeed, they are in all likelihood an indirect measure of motivation. This statement is predicated, however, upon the intelligence factor being within a normal range. Assuming intelligence to be at least normal, then, the writer believes that high school grades may well prove to be more dependent upon motivation than upon intelligence. Goodenough states,

That academic achievement is not substantially commensurate with native intellectual ability is one of standard credos of psychology.<sup>24</sup>

---

23. Cronbach, op. cit.

24. F. L. Goodenough, Developmental Psychology, (second edition; New York: Appleton-Century, 1945), Ch. 19.

The resolution of this question must be relegated to the future since we have no way of objectively measuring motivation at the present. It seems a reasonable hypothesis, however, that intelligence assumes a more predominant role at the college level, although correlation figures do not substantially uphold this line of reasoning. On the other hand, college students represent a more selective group and therefore one would expect a lower relationship index between the two variables. Thus when we combine high school standing and A.C.E. scores we may be combining four indices of measurement: (1) a sample of work habits, (2) motivation, (3) native intelligence, (4) a partial index of reading ability. Or it might be safer to say that none of these factors is measured in its entirety; rather each is measured partially, and all are inextricably interdependent. This would appear to be the only logical way to account for the above findings. More will be said of this in a later chapter.

#### The A.C.E. as a Differential Predictor

Since the A.C.E. is thought of as a placement and counseling tool as well as a predictor of general college success, a great deal of work has centered about the first-mentioned function. Various investigators have probed for possible differential qualities in the Q, L, or Total scores which would allow them to point the way to certain courses or careers.

MacPhail correlated Q and L scores with assorted college courses as well as with a group of "quantitative" and a group of "verbal" or linguistic courses.<sup>25</sup> The only significant correlation he could obtain from the entire study was between L scores and entire four year averages in verbal type subjects. This correlation was .50 with a Critical Ratio of 4.68. He cautions strongly against using the Q or L scores for guidance or counseling.

Salley and Weintraub compared high school standing, Regents Examinations and A.C.E. scores with grades made in four years of college.<sup>26</sup> They found the A.C.E. to be the least predictive of general college achievement, (.33), and in specific study areas of specialization it was also the least predictive. It is of great interest to the writer to note here that the investigators also state they found quite a few honor graduates who had been barely acceptable on all three measures at entrance, as well as quite a few borderline graduates who had been high scorers at entrance.

Super reports a summary which he made of research findings available, and, contrary to what one might expect, it

---

25. MacPhail, op. cit.

26. Ruth E. Salley and Ruth G. Weintraub, "Student Records of Entrance and Graduation," School and Society, 69:404-6, June 4, 1949.

showed L scores to be as good as Q scores for predicting success in the so-called "quantitative" courses of science and mathematics, and better than Q scores for prediction in English.<sup>27</sup> Also, it will be recalled from page 9 of this thesis that Wallace reported substantially the same thing. Anderson reports, as a result of an extensive study, portions of which were cited earlier, that he finds L scores to be better predictors of over-all success than Total scores.<sup>28</sup> He adds that Q scores and Wechsler Performance scores are not very dependable.

Hoerres,<sup>29</sup> Brown,<sup>30</sup> and Osborne, et. al.<sup>31</sup> all report substantially the same thing; L scores are more predictive in all fields than Q scores, (with some few exceptions) and they are particularly more predictive in linguistic or verbal type subjects. It is the concensus that Q should be used with caution, if at all.

---

27. Super, op. cit., p. 122.

28. Edward E. Anderson, et. al., op. cit.

29. Mary A. Hoerres, "A.C.E. Q, L, and T scores Vs. Freshman Grades at University of Wisconsin," Journal of Higher Education, 25:97, February, 1954.

30. Hugh S. Brown, "Differential Prediction by The A.C.E.," Journal of Educational Research, 44:116-21, October, 1950.

31. R. T. Osborne, Wilma B. Sanders, and J. E. Greene, "The Differential Prediction of College Marks by A.C.E. scores," American Psychologist, 4:286-7, July, 1949.

One final study turns what may prove to be quite an interesting light upon the subject. Smith and Triggs did a study on students who had high Q scores and low L scores.<sup>32</sup> They, like the others, found Q to be relatively a poor predictor of college success. Unlike the others, they decided to look for a reason. They conclude their report with the suggestion that perhaps the extent to which the Q abilities can be applied is limited by the level of L ability. This is an interesting suggestion. Certainly it appears to be a reasonable one.

#### Factors Which May Affect A.C.E. Scores and Validity

No review of research which reports as much fluctuation as this would be complete without an effort to discover the reason. Investigations have been made of several factors which were thought to be capable of influencing A.C.E. scores, and, in turn, the reliability and validity figures.

Admissions Policy. Travers and Wallace reported a study in which they tested the freshman classes of a dental college for two successive years.<sup>33</sup> With college grades as criteria they found a correlation of .10 with A.C.E. Total for the first

---

32. Smith and Triggs, op. cit.

33. R. M. W. Travers and W. L. Wallace, "Inconsistency in The Predictive Value of a Battery of Tests," Journal of Applied Psychology, 34:237-9, August 9, 1950.



group, but in the following year it was .48. They conclude that in a majority of cases admitted in the low correlation year the admissions officer used test scores alone, while in the year of better correlation he had used both test scores and previous grades. They suggest that the selection process is a very important one in controlling validity of predictive power for a battery of tests.

Age and Previous Military Status. The studies on this are not in complete agreement although there seems to be no really significant lack of accord. Owens and Owens studied 194 male veterans and found a correlation of .47 between A.C.E. scores and grades.<sup>34</sup> Age gave a slight positive R with grades, and length of time in service gave slight negative R. By taking all three into account they obtained an R of .57.

Pierson used age, sex, marital and military status Vs. Total A.C.E. in single and multiple correlations and reported that he found no single or combined factors significantly predominant.<sup>35</sup> Frederiksen and Schrader found the A.C.E. slightly more predictive for a group of veterans than for a group of non-veterans, but the difference was not too significant.<sup>36</sup>

---

34. William A. Owens and William A. Owens, Jr., "Some Factors in The Academic Superiority of Veteran Students," Journal of Educational Psychology, 40:497-502, 1949.

35. Rowland R. Pierson, "Age Vs. Academic Success in College Students," School and Society, 68:94-5, 1948.

36. Frederiksen and Schrader, op. cit.

Sex. Although females tend to score higher on some parts of the test and males higher on others, and there are other variations commensurate with type of college, geographical region, and the like,<sup>37</sup> sex appears to have little effect upon the predictability of college success. Osborne, Sanders and Greene report that on the L scores females predict somewhat better than males in certain courses, but this does not appear to be enough basis for expecting a consistent difference.<sup>38</sup>

Other Factors. Cronbach states that while the A.C.E. estimates probable college success as well as any other test available, he feels that slow-but-accurate workers are penalized in the score no matter how capable they might be.<sup>39</sup> This is because the score is calculated simply by totaling the correct items. Fusfeld says that academic success is not any more closely commensurate with college success because of in-college factors such as the time demands of extra-curricular activities, athletics, money earning and the like.<sup>40</sup> While this may be true to some degree, the writer disagrees with his

---

37. For more complete information the reader is referred to the Final Reports issued by the Educational Testing Service, op. cit.

38. Osborne, Sanders and Greene, op. cit.

39. Cronbach, op. cit., p. 168.

40. Irving S. Fusfeld, "One The A.C.E. Psychological Examination," School and Society, 70:117-18, August 20, 1949.

suggestion that these factors be ruled out in validating the test. At least, such a validation figure would not be very helpful for our purposes since we are counseling with and admitting students who are affected by these factors, and we wish to know how valid the test is for them. Such a study might be helpful for scientific research, however.

The Wheelers believe that we are testing a great deal more reading ability with the A.C.E. than intelligence.<sup>41</sup> They studied 1,681 university freshmen and found that the Total A.C.E. scores correlated .71 with total reading scores, (S.E. = .012). Comparing this with the findings of Anderson, (.45 correlation between the Wechsler-Bellevue and the A.C.E.),<sup>42</sup> one is inclined to agree. Thus we find some support for our supposition that reading ability is of primary importance in academic success.

## II. LITERATURE ON THE COOPERATIVE READING TEST

Not nearly the work has been reported on the Cooperative Reading Test as on the A.C.E. Possibly it is because this test is a sub-test of the Cooperative English Test, although it is a separate entity unto itself.

---

41. Lester R. Wheeler and Viola D. Wheeler, "The Relationship Between Reading and Intelligence Among University Freshmen," Journal of Educational Psychology, 40:230-8, April, 1949.

42. Anderson, et. al., op. cit.

### Reliability

There is little doubt that what we are measuring with the Cooperative Reading Test we are measuring with a good degree of consistency. In his review in the Third Mental Measurements Yearbook Robert M. Bear states that reliability coefficients as high as .90 have been reported for the test.<sup>43</sup> In the same volume J. B. Stroud says that it is well made and the best test of its kind available today.<sup>44</sup>

Mrs. Frances Gaver of the Educational Testing Service states that they have ascertained the reliability of Form RX of this test (the form with which we are concerned) in terms of standard errors of measurement at specified scaled score points.<sup>45</sup> The figures are shown below:

<u>Total Reading Score</u>	<u>S.E.</u>
50	2.5
70	3.0

According to Garrett,<sup>46</sup> standard error of measurement is one of the best methods of stating reliability and this S.E.

---

43. Buros, Third Mental Measurements Yearbook, op. cit., p.497.

44. Ibid, p. 498.

45. Letter, from Mrs. Frances Gaver, Assistant Editor, Cooperative Test Division, Educational Testing Service, Princeton, New Jersey, July 8, 1955.

46. Henry E. Garrett, Statistics in Psychology and Education, (Third edition, New York: Longmans, Green and Company, 1947).

indicates high reliability for the Cooperative Reading Test. Thus it would appear that the reliability of this test is as high as can reasonably be desired.

### Validity

No other area illustrates more clearly than reading that tests bearing the same name measure quite different behaviors...One author examined 24 different reading tests and found that between them they measured 48 different skills.<sup>47</sup>

There is indeed much difference of opinion about that which one is measuring, or should be measuring, with a reading test, and this somewhat complicates the method of validating it. For our purposes, and for want of a better one, we will have to accept marks in college courses as a criterion. This is based on the assumption that these marks reflect reading ability, which is a reasonable condition.

Buckton has done some rather extensive studies at Brooklyn College.<sup>48</sup> He correlated Total Reading scores with college marks for approximately 1,000 freshmen, and 1,000 sophmores, juniors and seniors who had had two or more courses in English while in college. Also he correlated the marks of those with

---

47. Cronbach, op. cit., p. 287.

48. LaVerne Buckton, "The Prediction of Student Success at Brooklyn College," Reported in a letter from Mrs. Frances Gaver, op. cit.

four or more courses in English and those who were English majors. His correlations ran in the low .30's with the exception of those for the English majors which ran from .35 to .40.

Humber, in 1942, tested a large number of seniors with the A.C.E. and the Cooperative Reading Test and correlated the results with grades made in their major fields.<sup>49</sup> He reported that the Cooperative Reading Test (as well as the other reading tests used) correlated better than the A.C.E. in practically every subject. He states that the difference between an "A" and a "C" effectively depends upon reading ability.

In summarizing the above we might say that the Cooperative Reading Test possesses good reliability and a fair degree of positive correlation with college grades in general, and that it has a tendency for differential prediction in favor of the humanities.

---

49. W. J. Humber, "The Relationship Between Reading Efficiency and Academic Success in Selected University Curricula," Journal of Educational Psychology. 35:17-26, January, 1944.

## CHAPTER III

### METHODS AND PROCEDURE

The general purpose of this study was to ascertain the degree of statistical relationship existing between college success and scores obtained on two tests included in a battery of tests administered to incoming freshmen at University of Richmond. The methods employed are outlined below.

#### I. CHOICE OF CRITERION

The writer had two main criteria to choose between: college marks, or success and failure as indicated by graduation from college vs. failure. The latter was chosen for several reasons.

First, had we chosen grades we would have been limited in the size of our sample, for quite a few left college in the first two semesters of attendance. Not only would this have cut the size of our sample drastically, but it would have meant that we could have used only a relatively short grading period, the limits of which would have been determined by the length of time-in-college common to all subjects. Certainly we could not have used a four year average with the same weight as a

one or two year average. Also, we found that the summer school grades were not included in the posted averages of students, and this factor would certainly distort the picture. Another important factor which militated against using grades was the apparent inconsistency in two graders' appraisal of the same student. It is a matter of common knowledge that teachers vary in the amounts of subjectivity and objectivity employed in testing and appraising the academic standing of their students. Even the qualities and work upon which they are graded vary from course to course.

On the other hand, if a student fails to pass the required number of hours necessary to justify his retention as a student in the college we can be assured that more than one professor has had to classify his work as below an acceptable level, and that his dismissal was a carefully considered action.

For another reason, this writer desired to try a different method of attack on the problem inasmuch as the grade criterion had been used over and over again, whereas the pass vs. fail method is apparently unique for this problem. It was hoped that results obtained by this method might prove to be more meaningful.

In order to have any significance, a grade criterion should certainly include two or more semester's grades; preferably, at least four, in order to allow the student time to adjust to the college situation. However, if this is done we



will miss a majority of the drop-outs and failures which occur before the start of the junior year. Thus our obtained figure of correlation is primarily applicable to those students who are not going to fail, and we are guilty of taking what is not a random sample representative of all college entrants, but a very select sample from which most failures have been eliminated. This figure would be useful, perhaps, in working with students who will succeed in college, but it has little meaning for the others.

Finally, it was hoped, by using pass vs. fail as a criterion, to obtain some figures which would be useful in working out a critical score for admissions to college.

## II. SUBJECTS

As subjects for this study it was decided to use the freshman class of 1949. This was desirable from several standpoints.

In the first place, by 1949 the great influx of veterans had dwindled nearly to the proportions in which we now have them, thereby giving us a sample more nearly normal. Secondly, this was the earliest year in which the program of testing was in effect at University of Richmond, and the writer wished to allow the subjects ample time to graduate or to become academic casualties. This date allowed them well over four

years since the time they entered college.

Also, choosing the earliest class possible gave more time for other workers to study the problem and to publish their results. This has proven very helpful.

The subjects, then, were chosen from a group of 258 freshmen who entered the University of Richmond in September of 1949. Of this group 87 had to be eliminated for the reasons indicated in Table I below.

TABLE I  
STUDENTS ELIMINATED FROM SUBJECT SAMPLE  
BY REASONS WHICH DISQUALIFIED THEM

Reason student not counted	Number students in category
Still in school	9
Transferred, not failing	23
Dropped out, unexplained, not failing	38
Dismissed, Honor Code violations	10
No A.C.E. score recorded	6
Not located in files	1
Total Students	87

These students obviously could not be counted as having "passed" or "failed" because they had done neither at the time of the study.

The "pass" group was chosen simply on the basis of its members having already received their degrees, or being candidates for degrees at the following convocation. The latter were chosen from advance programs of graduation. The pass group consisted of 92 students.

Membership in the "fail" group was determined by having left school one or more times as an academic casualty and not having returned as of the date of this study. Table II, below, will give the reader a clearer idea of casualty rate as the semesters progress through four years of college.

TABLE II  
NUMBER OF STUDENTS ELIMINATED UNDER AUTOMATIC RULING  
BY SEMESTERS, CLASS OF 1954

	Semester								Total
	1	2	3	4	5	6	7	8	8
No. Students Eliminated	32	15	7	12	3	6	1	3	79

It will be noted that the greatest rate of casualty is in the first two semesters with the average between the second and third semesters. The number who qualified as members of the "fail" group was 79.

### III. THE TESTS EMPLOYED

The two tests used in this study have already been mentioned. They were, (1) The American Council on Education Psychological Examination For College Freshmen, (2) the Cooperative Reading Test.

Both of these tests were scored and recorded upon IBM cards. The A.C.E. scores are reported as raw scores and are used throughout as such. This test yields a sub-score supposedly denoting quantitative reasoning ability called "Q" score, and a verbal ability sub-score reported as "L" score. Their total comprises the student's "T" score.

The reading test is a sub-test of the Cooperative English Test. It is composed of two parts, (1) a recognition vocabulary test with a separately reported score, (2) a reading comprehension test in which the student is scored for speed of comprehension and level of comprehension for the same performance. The latter two scores are reported separately. The three scores just mentioned are combined and reported as a Total Reading Score. All of these scores are scaled, and the Total Reading Score becomes a part of the Total English Score when the test is given as part of a battery. Only the Total Reading Scores were used in this study.

## IV. PROCEDURE

General

Having obtained IBM cards with the scores and names of each student, the writer, with the cooperation of the Dean's Office and the college registrar, checked the college records and obtained the information cited above. On the basis of this information the students were divided into the Pass, Fail, and Neither groups. The Neither group was rejected and frequency distribution tables were set up (one for the Pass group, one for the Fail group and one for a combination of both called the Total group) in each of the four score categories, (Q, L, T and Total Reading.). Modified and shortened forms of these tables appear in the Appendix.

From the data grouped into the above tables the following information was obtained for each score category, (1) the means of the Pass, Fail and Total groups, (2) the standard error of each mean, (3) the standard deviation of each distribution, and (4) the range of each score distribution. For sample calculations see the Appendix.

All mathematical calculations were checked twice to insure accuracy.

Using the data accumulated at this point, the writer then proceeded to calculate the Biserial Correlations for each score category as well as the standard error for each  $R_{bis}$ .

### Statistical Method Employed in Obtaining Correlations.

The method of correlation employed was that of obtaining the biserial R as outlined by McNemar,<sup>1</sup> and by Smith<sup>2</sup>. Inasmuch as the method used in this calculation is so well established and accepted, the mathematics upon which it is based need not be discussed here. The reader's attention is invited, however, to the discussions in the works cited above as well as to others cited in the Bibliography. Sample calculations are given in the Appendix.

While the statistical method is not in question, there is sometimes the question of whether its use is justified. Biserial correlations are usually employed when the experimenter wishes to obtain the relationship ratio between two variables, one of which is graduated in nature and yields an approximately normal distribution while the other is dichotomized. The basic assumption necessary before one may use the biserial R is that the dichotomized variable is actually continuous in nature and of fairly normal distribution; otherwise the point biserial correlation is called for. As McNemar points out "The tenability of the assumption of a continuous normally distributed variate

---

1. Quinn McNemar, Psychological Statistics, (New York: John Wiley and Sons, Inc., 1949) pp. 166-74.

2. G. Milton Smith, A Simplified Guide to Statistics for Psychology and Education, Revised edition; New York: Rinehart and Company, 1946.), pp. 81-84.

underlying the dichotomized trait must always be faced by the user of biserial R."<sup>3</sup>

In the case of this study the writer feels reasonably certain that the dichotomized groups (i.e. the Pass and Fail groups) do represent a continuous variable. In the first place, although perforce students are divided by a sharp passing-failing line, it is easily recognized that "pass" stands for everything from the barest acceptability to Phi Beta Kappa. There is little argument here. On the other hand, the Fail group most certainly must contain everything from a dismal failure to a near pass. One index of the continuity of the trait in this group is illustrated in Table II (above) where we see that some failed immediately, while it took others eight semesters.

In view of these facts the writer feels justified in using the biserial R rather than the point biserial R which is often less accurate.

---

3. McNemar, op. cit., p. 173.

---

## CHAPTER IV

### RESULTS

In this chapter the reader will find the results of this study reported for the most part in tabular form. This was felt to be more desirable because of the mathematical nature of the raw findings. Later in the chapter a summary and explanation of the data will be presented along with a critique of their agreement with published studies.

The test scores which are hereinafter represented by Q, L, T and Total Reading have been explained in the preceding chapter. In order to facilitate the reading of tables the following key of abbreviations and symbols is presented:

M	-	Mean
SE <sub>m</sub>	-	Standard Error of Obtained Mean
SD	-	Standard Deviation
R <sub>bis</sub>	-	Biserial Coefficient of Correlation
SE <sub>Rb</sub>	-	Standard Error of R <sub>bis</sub>
CRR <sub>b</sub>	-	Critical Ratio of R <sub>bis</sub>
R	-	Range of Scores
N	-	Number Subjects in Sample

Sample calculations and frequency distribution tables will be found in the appendix.



TABLE III  
 CONSOLIDATED SUMMARY OF DATA

	Pass Group	Fail Group	Total Group
- Q scores -			
M	47.61	39.95	44.07
SE <sub>m</sub>	1.03	1.18	.83
SD	9.82	10.45	10.81
R <sub>bis</sub>	-	-	.44
SER <sub>b</sub>	-	-	.081
CRR <sub>b</sub>	-	-	5.417
R	47.	56.	56.
- L scores -			
M	69.23	56.90	63.53
SE <sub>m</sub>	1.46	1.68	1.20
SD	13.92	14.88	15.63
R <sub>bis</sub>	-	-	.49
SER <sub>bis</sub>	-	-	.077
CRR <sub>b</sub>	-	-	6.379
R	54.	84.	84.
- T scores -			
M	116.50	96.95	107.44
SE <sub>m</sub>	2.27	2.61	1.87
SD	21.66	23.01	24.36
R <sub>bis</sub>	-	-	.50
SER <sub>b</sub>	-	-	.077
CRR <sub>b</sub>	-	-	8.327
R	97.	132.	132.
- Total Reading scores -			
M	57.73	49.11	53.75
SE <sub>m</sub>	2.9	2.92	.78
SD	9.4	8.09	9.82
R <sub>bis</sub>	-	-	.55
SER <sub>bis</sub>	-	-	.073
CRR <sub>b</sub>	-	-	7.542
R	44.	39.	47.
N	92	79	171

TABLE IV

CORRELATIONS OF A.C.E. SCORES Q, L, T,  
AND COOPERATIVE READING TEST TOTAL SCORES  
WITH COLLEGE SUCCESS AT UNIVERSITY OF RICHMOND\*

Test score	Rbis	SEgb
A.C.E. Q	.44	.081
A.C.E. L	.49	.077
A.C.E. T	.50	.077
TOTAL READ.	.55	.073

\* = Taken from Table III

It is interesting to note in Table IV (above) that Total Reading shows the highest correlation and the lowest standard error while just the reverse is true of Q.

TABLE V

COMPARISON OF MEAN SCORES FOR NATION,  
CLASS OF 1954\* AND STUDY GROUP

	Q scores		L scores		T scores		Total Read.	
	M	SD	M	SD	M	SD	M	SD
Nation <sup>1</sup>	40.0	11.8	62.0	17.7	102.0	27.1	54.7	10.5
U. of R. <sup>2</sup>	43.9		63.2		107.1		53.4	
Study Gp.	44.1	10.8	63.5	15.6	107.4	24.4	53.8	9.8

\* = Class of 1954 is Freshman Class of 1949

1. Educational Testing Service, Report on The 1949 College Freshman Testing Program, Appendix A, Table 1.

2. Educational Testing Service, Report to University of Richmond on The 1949 College Freshman Testing Program,

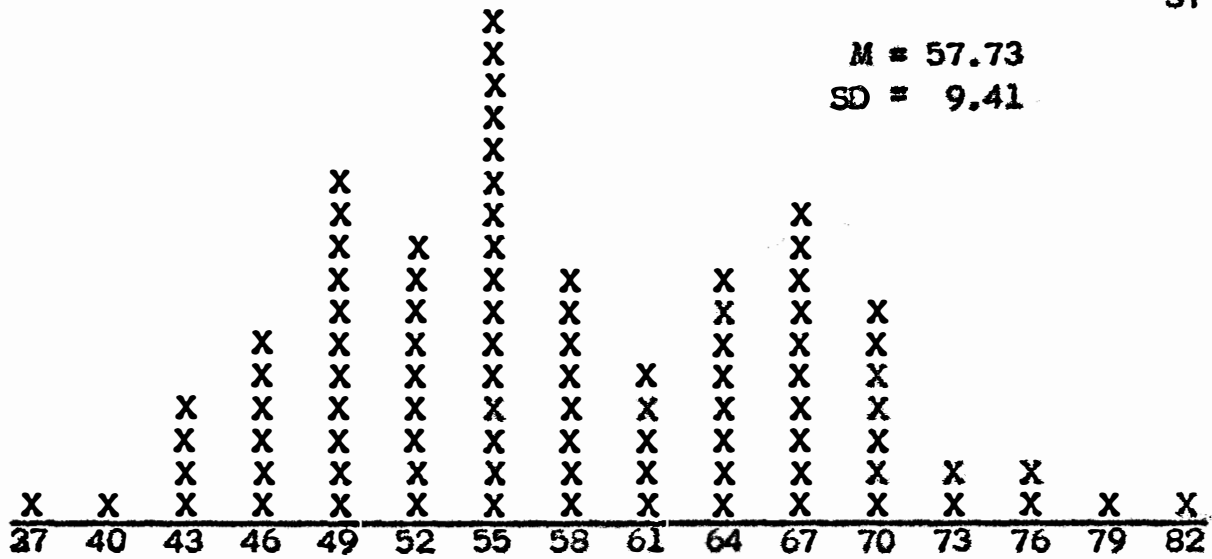


FIGURE 1

PASS GROUP: DISTRIBUTION OF SCORES ON COOP. READING TEST

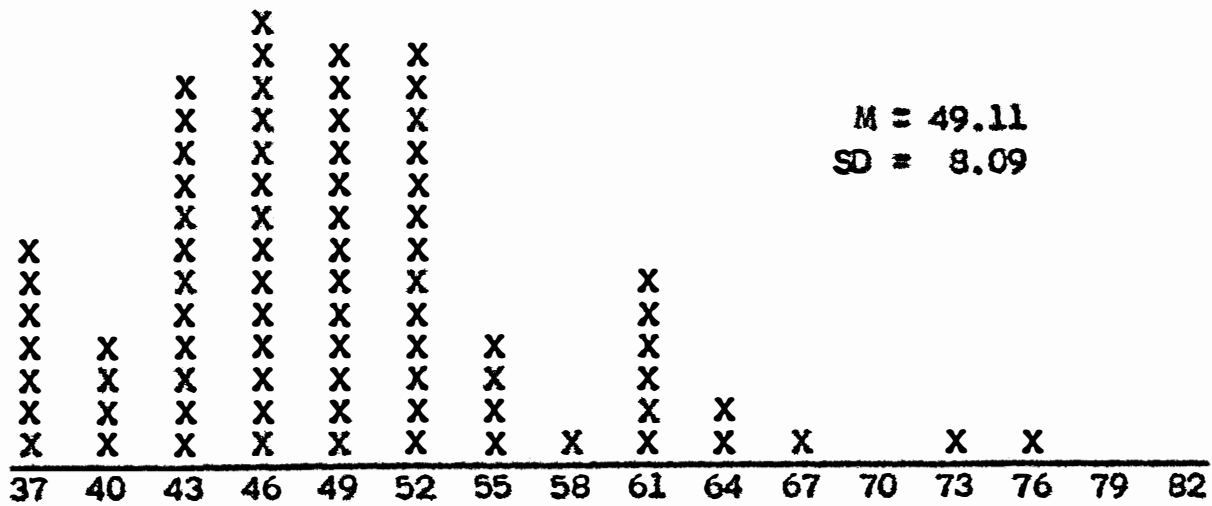


FIGURE 2

FAIL GROUP: DISTRIBUTION OF SCORES ON COOP. READING TEST



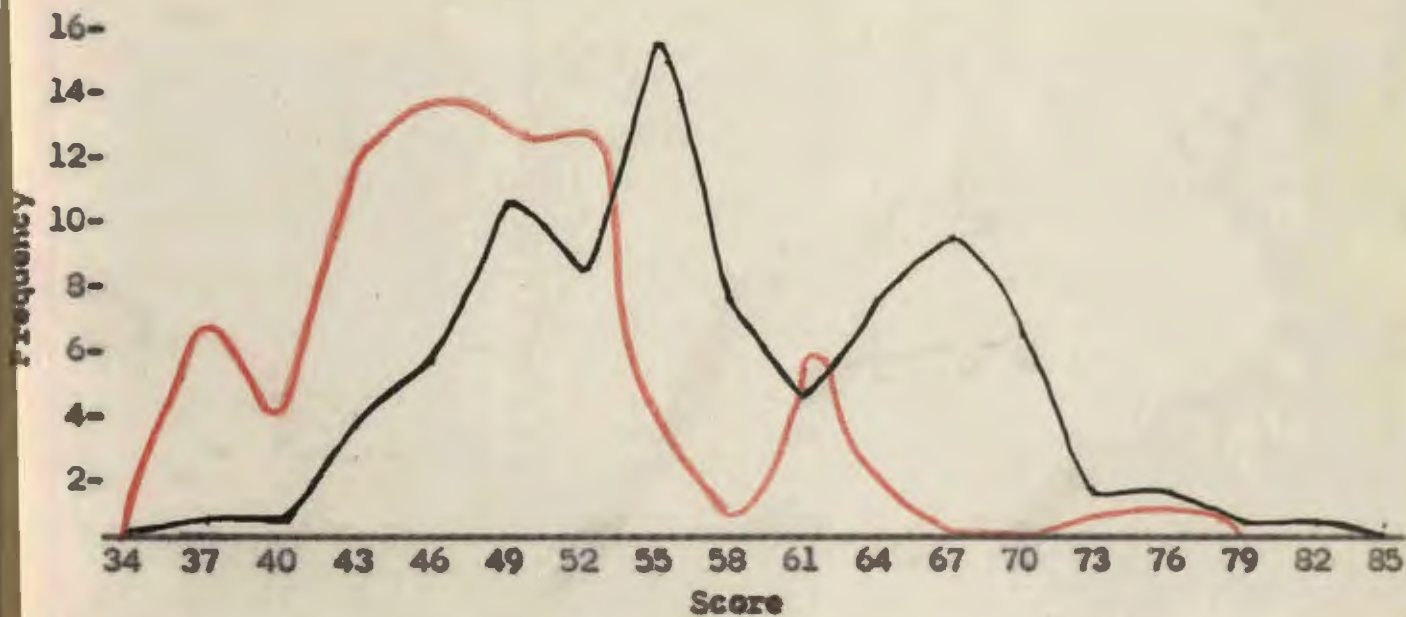


FIGURE 5

GRAPH OF SCORES OBTAINED ON COOP. READING TEST  
 BY PASS AND FAIL GROUPS  
 (FROM FIG. 1 AND FIG. 2)

— Pass group  
 — Fail group

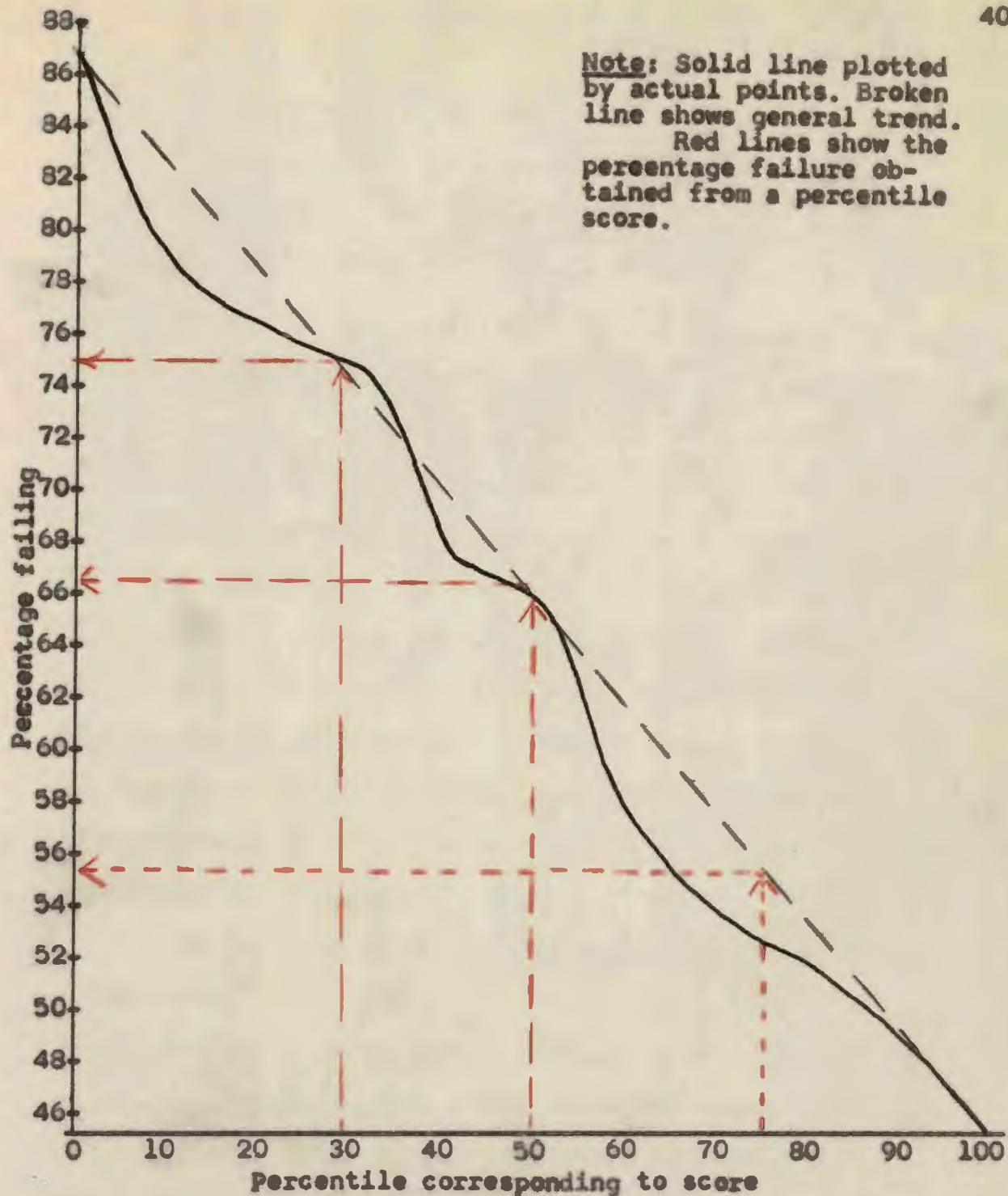


FIGURE 6

PERCENTAGE FAILURE RATE OF STUDENTS SCORING AT OR BELOW  
 A GIVEN PERCENTILE ON THE COOP. READING TEST.  
 PERCENTILES ARE THOSE FOR UNIVERSITY OF  
 RICHMOND, 1949

## I. SUMMARY OF RESULTS

From the results obtained it may be said that the Total Reading scores of the Cooperative Reading Tests are more closely related to college success in the sample studied than either the Q, L or Total scores on the A.C.E. This, of course, must be said with the reservation that this proved to be the case for the freshman class entering in 1949. On the other hand a look at Table III, page 35, will show such a low standard error of the mean that one may be reasonably certain that this result did not occur through chance. Also, the critical ratio of  $R_{bis}$  is comfortably high. The lowest reported is for Q scores (5.479) which again assures us that we may accept the results with a maximum degree of certainty. According to Smith<sup>3</sup> the chances are .00006, or less, in 100 that the obtained value could occur on the basis of chance variations in sampling.

The test score showing the least degree of positive relationship was Q as we might well have expected from a study of other investigators' findings. The best predictor on the A.C.E test turned out to be the T score, although it still did not equal the predictability of Total Reading, (see Table IV, page 36, for a comparison).

---

3. Smith, op. cit., p. 59.

In Figures 1, 2, 3, 4 on pages 37 and 38 the reader will see the distributions of scores made by Pass and Fail groups on the A.C.E. total and Total Reading. These represent fairly normal distribution curves for a sample of this size, (with the exception of that for the Fail group in Figure 2 which is slightly skewed toward the lower end.) This again lends support to the validity of our results. It is interesting to note the extreme ranges represented by these distributions, however. They range from the 1st to the 100th percentiles. There might be a suggestion here for a more carefully controlled admissions policy.

On the other hand, when these curves are plotted, Figure 5, page 39, we see that the Pass group exhibits the greatest range. A perusal of this figure will show that a large number of the Fail group exceeded the scores of many of the Pass group. This is in accord with the findings of Salley and Weintraub (see page 16 of this thesis) and again reminds us that test scores are still not the last word in academic achievement. Certainly those students scoring in the top percentiles of the A.C.E. and Cooperative Reading Test possess the wherewithal to make the grade if those scoring in the very lowest range can make it. In one particular case, IBM card number 20560, the student made the highest score of the study on the A.C.E. and almost the highest on the reading test;



yet he was a borderline case for four semesters and finally was eliminated under the automatic ruling. We know the false basis of generalizing on a few cases, but a glance at figure 5 shows that there are more than a few. While our correlation figures show that these are the exceptions one is still appalled at the great degree of overlap in this Figure and what it represents. In all likelihood a multiplicity of factors account for this discrepancy, among which motivation and work habits are the most outstanding.

In Figure 6, page 40, is a graph based on the Cooperative Reading Test in which the probability of failure is shown as a function of the percentile score made by a student. This graph was constructed by obtaining the cumulative frequency at each decile and calculating the percentage of failure for the students at, or below, these points. Percentile tables made up for University of Richmond by the Educational Testing Service were obtained from the Dean's Office, and the test scores converted.

If it were shown later that this table were fairly accurate for successive classes it would be a simple matter to construct a similar one for guidance and admissions use. As far as the class of 1954 is concerned it is accurate, of course, but the writer cautions against using it to calculate expectancy of academic success for any other class without

further research.

The graph is easily used. All that is necessary is to convert the student's Total Reading score into the equivalent percentile score by use of the proper tables, and then look up the percentage failure corresponding to this value on the graph

The total failure rate for the class of 1954 was 46.3% when we do not count the Neither group, and there appears to be no way that we can count it. Of the students that started only 30.3% actually failed. The other calculation is based on the assumption that those in the Pass and Fail groups at least stayed there and tried until they did one or the other, while the other group left before they did either.

## II. RESULTS COMPARED WITH THOSE PUBLISHED

In general, the results obtained in this study are substantiated by the findings of others. In fact, they are surprisingly comparable when we consider the extreme national variations reported by the Educational Testing Service (see Chapter II).

Table 5, page 36, gives a comparison of means for the nation, the entire class of 1954 and our sample. The reader will note that the means shown compare closely with those of this study, and that the standard deviations are close also.

It is interesting to note that all of our standard deviations are smaller than those of the national report, while our A.C.E. scores are higher. The one exception to this is in the Total Reading score where we see the study group is slightly higher than the total class but both are five percentile points below the national mean for male colleges. If these differences turn out to be statistically significant it would appear that we are better than average on the A.C.E. index but had quite a few poor readers in that class. The very close approximation of our findings to the national norms and to those for the entire class tends to reassure us that we did not get a biased sample and that our arithmetic is correct.

The above is about the best index of comparison available to us since this writer has found no other study using the same criteria. There are quite a few findings which support our own, however, if we consider grades fairly comparable as a criterion. Cronbach found about the same relative degrees of correlation for Q, L, and T, (.36, .38 and .41 respectively),<sup>4</sup> while Frederiksen found .47 for T. While lower correlations, and some few higher, have been reported our Q, L, and T scores have about the same relative predictability, (from lowest to highest in the order given), and it will be recalled from

---

4. All reference, unless it is indicated otherwise, refers back to earlier citations in Chapter II of this thesis. It is assumed that the reader has perused the thesis in the order of chapters.

the summary of findings in Chapter II of this paper that A.C.E. scores have generally been found to correlate in the .40's although the range reported runs from the low .30's to the mid .60's. Almost all studies have shown Q scores to be much lower than L or T scores in predictability. In fact, this writer can recall no study to the contrary.

In comparing results concerning the Cooperative Reading Test we are somewhat at a loss because of the paucity of comparable research. A good deal has been done on reading tests in general, but not too much has been reported by way of validations for this particular one. Buckton reported R's in the low .30's to .40's, but the best study this writer has found was Humber's in which he found that the Cooperative Reading Test correlated better than the A.C.E. in all subjects. The difficulty here is that no coefficients were reported. Even so, this corroborates our findings to some degree since this was exactly the relation we noted. In negative fashion it may be stated that there is little disagreement with the findings of this study.

In summation of the results of this study we offer that: our distributions are fairly normal, our standard errors low, our critical ratios high, our number of subjects adequate, and our comparisons with independent findings commensurate with normal expectancy. Although the study was not controlled

for sex, age, military experience and the like, we have cited independent research which attributes little or no weight to these factors.

In the light of the above, then, it appears quite reasonable to conclude that the findings of this study are valid and statistically sound.

## CHAPTER V

### SUMMARY AND CONCLUSIONS

#### I. SUMMARY

Beginning with the college year 1949 the University of Richmond has had in effect a program in which the successive freshman classes have been tested with a battery of tests supplied by the Educational Testing Service of Princeton, New Jersey. In light of the knowledge that injudicious use is often made of such tests when their limitations are not known, and also because it was felt that little was known about the actual efficiency of the testing program as it applied to this institution, it became the purpose of this study to investigate some aspects of a portion of the battery used.

Because of the position of relative importance accorded it the American Council on Education Psychological Examination for College Freshman was chosen as one of the tests. The other was chosen partially at the suggestion of a member of the Department of Psychology in the hope that it might lend support to the theory that reading was of fundamental importance to college success. This test was the Cooperative Reading test.

The total scores of both tests as well as the quantitative and linguistic sub-indices of the former were correlated with college success as measured by graduation or failure under the automatic ruling. Subjects chosen were members of the entering class of 1949. Of the original 258, 87 had to be discarded because they had not either graduated nor been dropped from school. 79 of the remainder classified as failures and 92 had graduated.

Biserial correlations were worked out with the result that a fair-to-good positive relation with college success was noted. Q scores yielded .44, L scores .49, T scores .50, while Total Reading yielded an even higher correlation of .55.

The writer interprets this to mean that, for the group studied at least, reading efficiency bears the highest degree of relationship to college success. From the magnitude of the error measurements found and reported, as well as from the size of the sample and the apparent agreement with independent findings the writer concluded the findings of this study to be as valid and as independent of chance results as could be hoped.

From the collateral search of pertinent literature the writer noted some other suggestions and facts which he would like to recapitulate here. There was information citing the high reliability of both tests as well as their sound construction and continued revision. On the other hand validity was

extremely variable. So much so that almost every author, as well as the publisher, admonished the user to work out the validity for his particular situation before placing a great deal of confidence in the results.

Several factors were cited which had been found to effect the validity of the A.C.E., among the most prominent of which was the college admissions policy; although age, sex, previous military experience and marital status were generally ruled out as insignificant. One author, from an analysis of the scoring system, felt that slow-but-accurate workers were penalized on the A.C.E. and there seems to be a good deal of logic in his criticism. Another author felt that the validity of the test was not higher because of time-demanding factors in college life.

The differential predictability of the test scores was found to be under some question. Most investigators agreed that the Q score appeared to be of slight value, if any, even in predicting success in mathematics and science courses. There was agreement that the L score seemed to have the greatest power of prediction, particularly in the linguistic courses, even more so than the T score. Again caution was cited in the use of these scores for guidance purposes.

Perhaps the most interesting finding in the literature was the obvious importance of high school standing in relation



to the prediction of college success. Studies showed it to be generally more predictive than any single test, or combination of tests. This index combined with the A.C.E. by multiple correlations yields the highest predictability available.

Total Reading and Total A.C.E. scores combined in a multiple correlation might show even higher predictability than either score taken singly. With the proper set of Beta weights, if this turns out to be the case, the guidance and admissions uses of these tests would be greatly enhanced. This particular correlation was not calculated because the data in this study was set up in frequency distribution tables (see Appendix) from which the proper figures were not available.

When compiling the data for this study it was also noted that while the mean scores for the A.C.E. at University of Richmond were slightly above national norms, the mean of the Total Reading scores was somewhat below the national mean.

In summation, we have "tested the test", or at least some aspects of it, and have found it as good as any available for the purpose although it is indeed limited.

## II. CONCLUSIONS

The writer finds that almost without knowing it the purpose of his study has transcended the original intention of validating a test for use in a particular population. In his

opinion this study has assumed the more important aspect of being a starting place for future studies in this area. It is not, in itself, as conclusive as the writer would have desired.

There are several things which have been established by this study, however. Reading ability, or at least that ability measured by a reading test, has been shown to be of considerable importance in the academic success of one class of students at this institution. Although it may reasonably be assumed that this applies in some degree to succeeding classes, the extent to which it applies can only be determined by actual investigations of these classes. This would be a highly worthwhile project if it resulted in establishing the consistency of reading importance.

Another conclusion which may be drawn from this study is the actual mathematical degree of correlation between the various predictive measures and college success. We have seen that the A.C.E. scores measure from the middle to the upper .40's while the Cooperative reading test exceeded the Total A.C.E. score's correlation by about 21%. This is of course predicated upon the fact that two correlations relate to each other as the ratio of their squares. Thus we know that while the A.C.E. measured as correlating positively and to a moderately good degree with college success, the Cooperative Reading

Test proved to be 21% again as good a predictor.

Perhaps even more important are the implications in this study. We have seen, in Chapter II, that the A.C.E. apparently is as predictive when administered in high school, even before the senior year, as when given in college. Also it has been demonstrated by other authors that high school standing combined with the A.C.E. score gives a very high predictive index. Thus it would not seem unreasonable to suggest that, if accuracy of prediction is desired, study be given to the implementation of such a program even if on a trial basis. If indications have been interpreted correctly by this writer the advantages of such a program are many fold, the greatest of which would be the additional time for logical planning made possible through accurate advance knowledge.

Another area in which this writer feels that further study may well prove rewarding is in the combination of the reading test score with high school standing to ascertain if an even higher correlation than those cited in Chapter II might be obtained. This would seem reasonable since reading scores appear to give a higher correlation than A.C.E. scores. Either of these studies could have a marked effect in the improvement of the admissions program at this institution.

Implicit in the findings of this study is the general importance of reading ability. We have seen earlier where one

author reported that reading ability effected the A.C.E. score. This writer has long felt that this may be one of the important factors behind the degree of correlation present in that test. The poor reader scores low because of his disability and this same disability causes him to get poor marks in school. The good reader does just the opposite and hence we find positive correlation. This of course is not held to be the only factor but it is almost unquestionably an important one. We have also seen that this study and others have shown reading ability to be closely allied with academic success, although reading ability and intelligence do not necessarily go hand in hand.<sup>1</sup> We find more and more often that studies are showing poor reading ability to be connected with college failure. There is a great deal more to be learned about this relationship, however, Further study in this area, particularly in connection with remedial reading, would be justified.

Another product of this investigation was the search for a critical or cut-off score which might be used in conjunction with the university admissions policy. The author did construct a graph (see Figure 6, page 40) showing the percentage of failure which could be expected of all those students scoring at or below a given percentile. Reading scores were picked as an index rather than A.C.E. scores because they exhibited a more consistent, straight-line relationship. While this graph

---

<sup>1</sup>I. R. W. Kilby, "Relation of A Remedial Reading Program to Scholastic Success in College". Journal Educational Psychology 36:573-34, Dec. 1945.

applies strictly to the class of 1954 and no other, it would not take a great deal in the way of subsequent investigation to determine if this failure percentage had a large enough degree of consistency from year to year to justify the use of such a device in prediction. This writer was not able to arrive at a definite critical score because of the overlap of scores cited in the previous chapter. For instance, if the University took only those scoring in the 50th percentile or better, (we are assuming for the moment that all classes will react similarly,) the failure rate would change from a calculated 46.3% (see page 44) to 29.6%. But while 68.5% of those turned down would have failed, 31.5% of those turned away would have made it. Thus any such score is clearly a matter of policy for the institution in question, for there is no critical score which would assure one of only turning down failures. There are still too many factors which we cannot yet measure. The foremost of these factors in this writer's opinion is motivation.

## BIBLIOGRAPHY

---

## BIBLIOGRAPHY

- Angoff, W. H. "Test Reliability and Effective Test Length," Psychometrika. 18:1-14, 1953.
- Anderson, Edward E., et al. "Wilson College Studies in Psychology: I. A Comparison of the Wechsler-Bellevue, Revised Stanford Binet, and American Council on Education Tests at the College Level," Journal of Psychology. 14:317-326, 1942.
- Bingham, W. VanDyke. Aptitudes and Aptitude Testing. New York: Harper and Brothers, 1937.
- Brown, Hugh S. "Differential Prediction by the A.C.E.," Journal of Educational Research, 44:116-121, October, 1950.
- Buckton, LaVerne. "The Prediction of Student Success at Brooklyn College," (Unpublished study obtainable at Educational Testing Service, Princeton, New Jersey.)
- Buros, Oscar K. (ed.). The 1940 Mental Measurements Yearbook. Highland Park, New Jersey: The Mental Measurements Yearbook, 1941.
- Buros, Oscar K. (ed.). The Third Mental Measurements Yearbook. New Brunswick: Rutgers University Press, 1949.
- Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Brothers, 1949.
- Educational Testing Service. Final Report of The 1952 National Freshman Testing Program. (Cooperative Test Division, Educational Testing Service Report.)
- Educational Testing Service. Report on The 1949 College Freshman Testing Program. (Cooperative Test Division, Educational Testing Service Report.)
- Educational Testing Service. Report to University of Richmond on The 1949 College Freshman Testing Program. (Cooperative Test Division, Educational Testing Service Report.)
- Educational Testing Service. "Summary of Selected Research on the Validity of The American Council Psychological Examination for College Freshmen as a Predictor of College Grades," Research Memorandum. (Mimeographed.), 1954.

- Frederiksen, Norman, and Schrader, W. B. "The A.C.E. Psychological Examination and High School Standing as Predictors of College Success," Journal of Applied Psychology, 36:261-265, August, 1954.
- Fusfeld, Irving S. "On The A.C.E. Psychological Examination," School and Society, 70:117-118, August, 1949.
- Garrett, Henry E. Statistics in Psychology and Education. Third edition. New York: Lingmans, Greene and Company, 1947.
- Goodenough, F. L. Developmental Psychology. Second edition. New York: Appleton-Century, 1945.
- Guilford, J. P. Psychometric Methods. Second Edition. New York: McGraw-Hill Book Company, 1954.
- Gulliksen, Harold. Theory of Mental Tests. New York: John Wiley and Sons, 1950.
- Hoerres, Mary A. "A.C.E. Q, L, and T Scores Vs. Freshman Grades at University of Wisconsin," Journal of Higher Education, 25:97, February, 1954.
- Humber, W. J. "The Relationship Between Reading Efficiency and Academic Success in Selected University Curricula," Journal of Educational Psychology, 35:17-26, January, 1944.
- Kilby, R. W. "Relation of a Remedial Reading Program to Scholastic Success in College," Journal of Educational Psychology, December, 36:513-34, 1945.
- Larsen, Harold D. (comp.). Rinehart Mathematical Tables. New York: Rinehart and Company, 1948.
- MacPhail, Andrew H. "Q and L Scores on the A.C.E. Psychological Examination," School and Society, 56:248-251, September 19, 1942.
- McNemar, Quinn. Psychological Statistics. New York: John Wiley and Sons, Inc., 1949.
- Mursell, James L. Psychological Testing. New York: Longmans, Green and Company, 1947.



- Osborne, R. T., Sanders, Wilma B., and Greene, J. B. "The Differential Prediction of College Marks By A.C.E. Scores," American Psychologist, 4:286-287, July, 1949.
- Owens, William A. and Owens, William A., Jr. "Some Factors in The Academic Superiority of Veteran Students," Journal of Educational Psychology, 40:499-502, 1949.
- Pierson, Rowland R. "Age Versus Academic Success in College Students," School and Society, 68:94-95, 1948.
- Remmers, H. H., Elliott, D. N. , and Gage, N. L. "Curricular Differences in Predicting Scholastic Achievement; Applications to Counseling," Journal of Educational Psychology, 40:385-394, November, 1949.
- Salley, Ruth E. and Weintraub, Ruth G. "Student Records of Entrance and Graduation," School and Society, 69:404-406, June 4, 1949.
- Samenfeld, Herbert W. "Predicting College Achievement," Journal of Higher Education, 24:432-433, November, 1955.
- Smith, D. D. and Triggs, Frances O. "Educational Successes and Failures of Students with High Q and Low L Scores on the A.C.E. Examination," American Psychologist, 5:353-354, July, 1950.
- Smith, D. D. and Triggs, Frances O. "Educational Successes and Failures of Students with High Q and Low L Scores on the A.C.E. Examination," (a follow-up on above article) American Psychologist, 6:295-296, July, 1951.
- Smith, G. Milton. A Simplified Guide to Statistics for Psychology and Education. Revised edition. New York: Rinehart and Company, 1950.
- Super, Donald E. Appraising Vocational Fitness by Means of Psychological Tests. New York: Harper and Brothers, 1949.
- Travers, R. M. W. and Wallace, W. L. "Inconsistency in The Predictive Value of a Battery of Tests," Journal of Applied Psychology, 34:237-239, August, 1950.
- Traxler, A. E. "Correlation Between Two Tests of Academic Aptitude," School and Society, 61:383-384, 1945.

- Wallace, W. L. "Differential Predictive Value of The A.C.E. Psychological Examination," School and Society, 70:23-25, July 9, 1949.
- Weintraub, R. G. and Salley, R. E. "Graduation Prospects of Entering Freshmen," Journal of Educational Research, 39:116-126, 1945.
- Wheeler, Lester R. and Wheeler, Viola D. "The Relationship Between Reading and Intelligence Among University Freshmen," Journal of Educational Psychology, 40:230-238, April, 1949.

**APPENDIX**

**—**

TABLE VI

## FREQUENCY DISTRIBUTION TABLE FOR TOTAL A.C.E. SCORES

Score	X	F <sub>p</sub>	F <sub>f</sub>	F <sub>t</sub>	FX	D	FD	FD <sup>2</sup>
162-168	165	0	1	1	165	57.56	57.56	3313.15
155-161	158	2	1	3	474	50.56	151.68	7666.94
148-154	151	3	3	6	906	43.56	261.36	11384.84
141-147	144	8	0	8	1152	36.56	292.48	10693.07
134-140	137	9	2	11	1507	29.56	325.16	9611.73
127-133	130	16	1	17	2210	22.56	383.52	8652.21
120-126	123	7	2	9	1107	15.56	140.04	2179.02
113-119	116	5	3	8	928	8.56	68.48	586.19
106-112	109	12	6	18	1962	1.56	28.08	43.80
99-105	102	5	17	22	2244	5.44	119.68	651.06
92- 98	95	15	11	26	2470	12.44	323.44	4023.59
85- 91	88	4	7	11	968	19.44	213.84	4157.03
78- 84	81	3	11	14	1134	26.44	370.16	9787.03
71- 77	74	1	8	9	666	33.44	300.96	10064.10
64- 70	67	1	3	4	268	40.44	161.76	6541.57
57- 63	60	1	1	2	120	47.44	94.88	4501.11
50- 56	53	0	1	1	53	54.44	54.44	2963.71
43- 49	46	0	0	0	0	61.44	0	0
36- 42	39	0	1	1	39	68.44	68.44	4684.03

Totals - 92 79 171 18373

 $N_p N_f N_t \sum FX$ 

101506.20

 $\sum FD^2$

TABLE VII  
 FREQUENCY DISTRIBUTION TABLE FOR L SCORES

Score	X	F <sub>p</sub>	F <sub>f</sub>	F <sub>t</sub>	D	FD	FD <sup>2</sup>
105-109	107	0	1	1	43.47	43.47	1889.64
100-104	102	0	0	0	38.47	0	0
95- 99	97	1	1	2	33.47	66.94	2240.48
90- 94	92	7	1	8	28.47	227.76	6484.33
85- 89	87	8	2	10	23.47	234.70	5508.41
80- 84	82	9	1	10	18.47	184.70	3411.41
75- 79	77	9	2	11	13.47	148.17	1995.85
70- 74	72	10	6	16	8.47	135.52	1147.85
65- 69	67	13	2	15	3.47	52.05	180.61
60- 64	62	8	16	24	1.53	36.72	56.18
55- 59	57	11	9	20	6.53	130.60	852.82
50- 54	52	10	14	24	11.53	276.72	3190.58
45- 49	47	3	9	12	16.53	198.36	3278.89
40- 44	42	3	8	11	21.53	236.83	5098.95
35- 39	37	0	6	3	26.53	79.59	2111.52
30- 34	32	0	3	3	31.53	94.59	2982.42
25- 29	27	0	1	1	36.53	36.53	1334.44
		92	79	171			41764.38
		↓	↓	↓			= ∑ FD <sup>2</sup>
		N <sub>P</sub>	N <sub>F</sub>	N <sub>T</sub>			

TABLE VIII  
FREQUENCY DISTRIBUTION TABLE FOR Q SCORES

Score	X	F <sub>p</sub>	F <sub>f</sub>	F <sub>t</sub>	D	FD	FD <sup>2</sup>
65 - 67	66	4	1	5	21.93	109.65	2404.62
62 - 64	63	3	3	6	18.93	113.58	2150.07
59 - 61	60	6	2	8	15.93	127.44	2030.12
56 - 58	57	6	2	8	12.93	103.44	1337.48
53 - 55	54	10	2	12	9.93	119.16	1183.26
50 - 52	51	10	0	10	6.93	69.30	480.25
47 - 49	48	13	5	18	3.93	70.74	278.01
44 - 46	45	10	7	17	.93	15.81	14.70
41 - 43	42	12	15	27	2.07	55.89	115.69
38 - 40	39	4	10	14	5.07	70.98	359.87
35 - 37	36	6	11	17	8.07	137.19	1107.12
29 - 31	30	0	7	7	14.07	98.49	1385.75
26 - 28	27	2	3	5	17.07	85.35	1456.93
23 - 25	24	0	0	0	20.07	0	0
20 - 22	21	1	2	3	23.07	69.21	1596.68
17 - 19	18	1	0	1	26.07	26.07	679.65
14 - 16	15	0	0	0	29.07	0	0
11 - 13	12	0	2	2	32.07	64.14	2056.97

92 79 171  
|   |   |  
N<sub>P</sub> N<sub>F</sub> N<sub>T</sub>

$\Sigma FD^2 = 19985.16$

TABLE IX  
 FREQUENCY DISTRIBUTION FOR TOTAL READING SCORES

Score	X	F <sub>p</sub>	F <sub>f</sub>	F <sub>t</sub>	D	FD	FD <sup>2</sup>
81 - 83	82	1	0	1	28.25	28.25	798.06
78 - 80	79	1	0	1	25.25	25.25	637.56
75 - 77	76	2	1	3	22.25	66.75	1485.19
72 - 74	73	2	1	3	19.25	57.75	1111.69
69 - 71	70	7	0	7	16.25	113.75	1848.44
66 - 68	67	10	1	11	13.25	145.75	1931.19
63 - 65	64	8	2	10	10.25	102.50	1050.63
60 - 62	61	5	6	11	7.25	79.75	578.19
57 - 59	58	8	1	9	4.25	38.25	162.56
54 - 56	55	16	4	20	1.25	25.00	31.25
51 - 53	52	9	13	22	1.75	38.50	67.38
48 - 50	49	11	13	24	4.75	114.00	541.50
45 - 47	46	6	14	20	7.75	155.00	1201.25
42 - 44	43	4	12	16	10.75	172.00	1849.00
39 - 41	40	1	4	5	13.75	68.75	945.31
36 - 38	37	1	7	8	16.75	134.00	2244.50

92 79 171  
 $N_P$   $N_F$   $N_T$

16483.70 =  $\sum FD^2$

TABLE X

66

SAMPLE CALCULATION OF BISERIAL COEFFICIENT OF CORRELATION  
BASED ON DATA FROM TABLE VI

$$1. M_T = \frac{\sum F_T X}{N_T} = \frac{18373}{171} = 107.44$$

$$2. SD_T = \sqrt{\frac{\sum F_T D^2}{N_T}} = \sqrt{\frac{101506.}{171}} = \sqrt{593.60} = 24.36$$

$$3. SE_M = \frac{SD}{\sqrt{N_T - 1}} = \frac{24.36}{\sqrt{171 - 1}} = \frac{24.36}{\sqrt{170}} = \frac{24.36}{13.04} = 1.87$$

$$4. R_{Bis} = \frac{(M_P - M_F) \times \frac{PQ}{Z}}{SD_T}$$

$$= \frac{(116.50 - 96.95) \times \frac{(.538 \times .462)}{.397}}{24.36}$$

$$= \frac{19.55}{24.36} \times \frac{.24856}{.397} = \frac{4.85926980}{9.6709200} = .50246$$

$$= .50$$

$$5. SE_{R_{Bis}} = \frac{\frac{\sqrt{PQ}}{Z} - R_{Bis}^2}{\sqrt{N_T}} = \frac{\frac{\sqrt{.24856}}{.397} - (.50)^2}{\sqrt{171}} = \frac{1.0030}{13.08} = .07668$$

$$= .077$$

$$6. CR_{R_{Bis}} = \frac{R_{Bis}}{SE_{R_{Bis}}} = \frac{.5024}{.0766} = 6.5587 = 6.559$$

Note: The subscript "t" stands for the total group, and "p" and "f" denote passing and failing groups.



Walter Nelson Taylor was born in Harrisonburg, Virginia on May 26, 1924. He attended public schools there and in Hanover County. During his senior year at high school, where he was class president, he enlisted in the Army Air Corps and served there in various capacities ranging from Radio-Operator-Gunner to Instructor until his discharge in 1946.

The following September he entered Randolph Macon College where he pursued a pre-medical course and received his Bachelors Degree in General Science in June of 1949. This same month he was married to Mae Justin Derieux of Urbanna.

The following fall he entered the School of Medicine at Medical College of Virginia where he studied until January, by which time he had decided that his real interests lay in other fields and resigned as a student in good standing. Later in the year he accepted a position teaching a combination fifth and sixth grade at Beaverdam School and subsequently decided to remain in the field of education. The following year he taught Math and Science at Ashland High School, and in the next year became principal of Henry Clay Elementary School in Ashland, Virginia where the present writing finds him.