1952

A study of the value of the differential aptitude tests for predicting success in ninth grade general science

BOSTON UNIVERSITY

SCHOOL OF EDUCATION

Thesis

A STUDY OF THE VALUE OF THE DIFFERENTIAL APTITUDE TESTS FOR PREDICTING SUCCESS IN NINTH GRADE GENERAL SCIENCE

Submitted by

George W. Angell, Jr.

(B.S. in Chem., Tufts, 1944)

In Partial Fulfillment of Requirements for

the Degree of Master of Education

1952

First Reader: William C. Kvaraceus

Professor of Education

Second Reader: James F. Baker

Assistant Professor of Education

TABLE OF CONTENTS

LIST OF TABLES

Table | Page

Table | Page

Table | Page

CHAPTER I

NATURE AND SCOPE OF PROBLEM

Statement of the Problem

This study is an attempt to determine statistically the value of the Differential Aptitude Tests in predicting success or failure in ninth grade English and general science in the Worcester and Newton junior high schools. The study is particularly pointed toward the field of general science.

Scope of the Study

The study is done on 146 ninth grade students, 84 girls and 62 boys, in the Worcester and Newton junior high schools. Near the close of the 1950 - 1951 school year, those students who had taken general science were given the Read General Science Achievement Test. Teachers' marks were obtained for ninth grade English and general science, and correlations were run with Differential Aptitude Test raw scores (total battery and the 8 separate sub-tests), in order to ascertain which factor or factors predicted success in each subject. Correlations were also run between Differential Aptitude Test raw scores and Read General Science Achievement Test scores, both to provide a check on the consistency of mark-

ing by the teachers, and to provide a common basis for the relative ranking of all cases in at least one subject.

Justification of the Study

Need for prognosis.-- Guidance workers, teachers, and administrative personnel are continually seeking better means of prognosis for educational and vocational purposes. The Differential Aptitude Tests attempt to provide, at least partially, some basis for reasonably accurate educational prognosis. They are not final answers, nor are they one hundred percent accurate. Indeed, the authors, Seashore, Bennett, and Wesman, point out the need for further validation studies.[1/]

Eurich and Cain[2/] stress the need for adequate objective measures of abilities and interests and say:

> ....newer practices involve a conception of an educational system that will provide for a maximum development of each child in terms of a given body of experiences and capacities. This purpose cannot be accomplished without adequate techniques for determining the variety of interests and capacities that give promise for the future. In the older programs it was possible to set up a course of study in terms of subject matter without much reference to the child being taught. The newer programs are wholly dependent upon the extent to which it is possible to look into the future to ascertain which abilities are worth developing. In terms of

1/G. K. Bennett, H. G. Seashore, and A. G. Wesman, Differential Aptitude Tests Manual, The Psychological Corporation, New York, 1947, p. E-1

2/W. S. Monroe, "Prognosis", Encyclopedia of Educational Research, MacMillan Company, New York, 1941, p. 878.

> the trends in American education, educational prognosis is essential to the point of determining the nature of the educational program for each child.

Need for research on multiple aptitude tests.-- Essentially, the multiple aptitude test stems from the conception of any aptitude as a combination of certain specific abilities thought to be necessary for success in a given area. If these specific abilities can be measured separately, or nearly so, then the relative weight of each measure for any given aptitude can be determined by use of the proper criteria. Measures of separate and specific abilities may then be used for accurate prediction of accomplishment in the criteria. This, simply, is the theory underlying the Differential Aptitude Tests.

The need for tests developed along these lines has been pointed out by a number of writers in the past twenty-five years. One of these is C. L. Hull[1/], who conducted one of the first analyses of trait differences. This analysis was based on thirty-five different tests given to one hundred and seven first-year high school students, and encompassed various separate trait measures. He enters a plea for further research with the words:

> ....What would it mean in terms of increased social efficiency and personal happiness if a system of voca-

1/C. L. Hull, "Variability in Amount of Different Traits Possessed by the Individual", Journal of Educational Psychology, Vol. 18, No. 2, February, 1927, pp. 97 - 106.

> tional prognosis could be established which would raise the average vocational choice from around the 50% level /of the individual's potentialities/ where it probably now stands to the upper 5% of individual potentialities? ....This would probably represent an advance comparable to anything which has taken place in modern times.

Cottle[1] praises the Differential Aptitude Tests, saying that they fill a long-felt need, but hastens to warn that the validation process is incomplete. He feels that the tests themselves are sound, but that adequate statistical proof of the validity is lacking, and that no final conclusions should be drawn on the basis of the presently existing evidence.

Yagjian[2], in a validation study of the Differential Aptitude Tests in relation to first-year high school bookkeeping, suggests that further validation studies of the Differential Aptitude Tests are needed in a variety of school subjects. She particularly notes that larger numbers of cases are needed to provide adequate data from which to draw generalizations as to the predictive value of the tests.

(

1/W. G. Cottle, "The Differential Aptitude Tests: Some Comments", Occupations, Vol. 26, No. 6, March, 1948, pp. 344 - 345.

2/C. Yagjian, A Validation Study of the Differential Aptitude Tests in Relation to First Year High School Vocational Bookkeeping in Two Communities, Unpublished Ed. M. Thesis, Boston University School of Education, 1950, p. 72.

Definition of Aptitude

There are many definitions of aptitude, most of which are too generalized for the purposes of this study. According to Crawford[1], aptitude is the ...."power to acquire /an/ ....ability through proper training and study". Good[2] defines aptitude as a ...."pronounced innate capacity for or ability in a given line of endeavor, such as a particular art, school subject, or vocation."

These definitions, while concise, are not specific enough to make clear the many aspects of aptitude which the term presents to the psychologist and guidance person. Bingham[3] has defined and redefined the word in several aspects, and has further defined the function of aptitude testing. He says:

> Aptitude, then, is a condition symptomatic of a person's relative fitness, of which one essential aspect is his readiness to acquire proficiency - his potential ability - and another is his readiness to develop an interest in exercising that ability. Anyone who has come to a clear realization of his capacities, his informed interests, and the nature of the occupations he is considering, has achieved some basis for appraising the relative strength of his aptitudes...

1/A. B. Crawford, "What About All These Tests?", Occupations, Vol. 12, No. 8, April, 1934, pp. 13 - 18.

2/C. V. Good, Dictionary of Education, McGraw-Hill Book Company, New York, 1925, p. 28.

3/W. V. Bingham, Aptitudes and Aptitude Testing, Harper & Brothers, New York, 1937, pp. 18 - 33.

..Aptitudes are, then, aspects of personality....[an] aptitude is a present condition, yes, but with a forward reference. It is a condition or set of characteristics regarded as symptomatic, indicative of potentialities. It should nevertheless be apparent that in measuring a person's aptitudes we are not undertaking to place a yardstick against mysterious intangibles. Neither do we attempt the legerdemain of measuring something which does not exist, namely, future accomplishments. A test of aptitude samples certain abilities and characteristics of the individual as he is today. It helps to find out what he can do now and how well he can do it. The responses he makes under specified conditions are ascertained - specimens of his performance when motivated in prescribed ways. By such means, data are secured as to what the person actually does under the circumstances imposed by the test. His behavior is measured. From these symptoms, any estimate of his future possibilities of accomplishment is an inference.

This is the gist of the theory of aptitude testing. Measure selected samples of a person's behavior, and then, by reference to the facts as to what others who have been tested have done subsequently, compute the probabilities that he, too, will behave in a certain manner.

....An individual's potentialities are not all equally strong....Individuals differ from one another in their potentialities....Many of these differences are relatively stable.

....The theory of aptitude implies that a person's potentialities are fairly stable....[it] must not assume that each of the factors determining a person's traits is constant in the sense that it will not alter, more or less, with time. We must, however, assume - if the concept of educational and vocational aptitudes is to have any meaning at all - that the changes which undoubtedly do take place in the relative potency of these factors are seldom sudden, and that they occur within limits which can often be ascertained in advance.

According to Wesman[1], an aptitude may be regarded as "a capacity to learn". He goes on to define the measurement

1/A. G. Wesman, "What Is An Aptitude?", Test Service Bulletin, No. 36, August, 1948, The Psychological Corporation, New York, p. 1.

of aptitude as:

> ....the assessment of knowledge, skill, and any other characteristics which serve to predict learning success. Sometimes background data which reveal experience, interest or personality characteristics may be equally fruitful sources of prediction.
> ....[an aptitude test] is any test which is used for the preducation of some type of learning. Its validity as an aptitude test depends on the extent to which it predicts successfully.

While Traxler[1] and Mursell[2] define the term aptitude approximately as does Bingham[3], Mursell cautions that "....we must avoid thinking of aptitudes as faculties, mental entities. Rather, they must be considered as dynamic trends of the whole personality".

Insofar as the Differential Aptitude Tests are concerned, a synthesis of the above definitions might be more concise and yet sufficiently descriptive. An aptitude, for the purposes of this study, may be defined as a combination of specific abilities necessary to the acquisition, through proper training, of some specific skill.

Procedure

The first portion of this study consisted of a thorough search of the literature to find information pertinent to

---

1/A. E. Traxler, "Evaluation of Aptitude and Achievement in a Guidance Program"; Educational and Psychological Measurement, Vol. 6, No. 1, Spring, 1946, pp. 3 - 16.

2/J. L. Mursell, Psychological Testing, Longmans, Green & Company, New York, 1950, p. 225.

3/op. cit., p. 20.

the theory and use of multiple aptitude tests in general, and the Differential Aptitude Tests in particular. Chapter II of this paper deals with prior research and theory underlying the concept of multiple factors composing specific aptitudes, prior work on multiple-factor aptitude measures, and research on the Differential Aptitude Tests themselves.

The next step was to obtain Differential Aptitude Test scores, for the total battery and for each sub-test, from the permanent record cards of the students in the Newton and Worcester junior high schools who had taken the Differential Aptitude Tests in the Fall of 1950 as a part of their usual testing program.

Once the results of the Differential Aptitude Tests battery had been obtained, it was necessary to obtain the criteria by which to judge the effectiveness of the instrument in question. This was done by obtaining the final teachers' marks in English and general science for each student. For purposes of checking on the consistency of teachers' marks and also to provide a basis for relative ranking of all cases in at least one subject, the Read General Science Achievement Test was used as a second criterion.

Finally, the predictor was correlated with the criteria, and conclusions were drawn therefrom.

The various measures of central tendency, i.e., mean, standard error of the mean, and standard deviation were cal-

culated for the total battery and for each sub-test of the Differential Aptitude Tests, for teachers' marks in each ninth grade subject, and for the Read General Science Achievement Test.

The results of the predictor and the criteria were correlated by means of the Pearson Product-Moment equation once it was determined by inspection that a linear relationship existed. Correlation coefficients were calculated for the following:

1. Between various sub-tests of the Differential Aptitude Test.
2. Between teachers' marks in English and general science and the Differential Aptitude Tests (total battery and individual sub-tests).
3. Between raw scores on the Read General Science Achievement Test and the Differential Aptitude Test raw scores (total battery and individual sub-tests).

Since the function of statistics is not only to describe data, but also to tell how well the data is described, the standard error of r has been calculated. Multiple correlations have also been calculated.

Re-statement of Problem

This study is concerned with further statistical validation of the Differential Aptitude Tests in prediction of

success or failure in ninth grade English and general science. The study is based on 84 girls and 62 boys from the junior high schools of Worcester and Newton.

CHAPTER II

REVIEW OF RELATED RESEARCH

Theory of Multiple Aptitude Tests

The theory underlying all multiple aptitude tests is based on the theories of Spearman, Thurstone, and Kelley. Although the original theories proposed by these men were quite different, they have been so revised that today there is little to distinguish between them. Spearman and Kelley exhaustively analyzed correlation tables, and Thurstone made a comprehensive factor analysis of a large battery of tests. The general conclusion was that intelligance was not a generalized thing, but a complex of many separate factors.

Hull and Limp[1/] have suggested that while much effort and attention has been devoted to one phase of differential psychology, individual differences, very little attention has been paid to the amount of difference between traits, abilities, and aptitudes within the individual. They have concluded that the chief value of differential psychology should lie in forecasting differences in extent of traits, abilities, and aptitudes of the individual.

1/C. L. Hull and C. E. Limp, "The Differentiation of the Aptitudes by Means of Test Batteries", Journal of Educational Psychology, Vol. 16, No. 2, February, 1925, pp. 73 - 88.

Hull and Limp organized aptitude test batteries using shorthand, typewriting, English and algebra marks as criteria. Forecasting efficiency ranged from 2.4% in typing to 17.1% in English. The test batteries correlated between .51 and .74 with the criteria, yet the inter-correlations were only between .17 and .50. This illustrated, the authors felt, that it was infinitely more difficult to distinguish between the amounts of various aptitudes of an individual than to distinguish between amounts of a single aptitude in different persons. They conclude that

> ....the study does show that these probably are real differences in the aptitude of individuals for such similar activities as learning the subjects taught in high school. The study suggests that a scientific vocational guidance may be realized by making forecasts on numerous type vocations from many test batteries made up from various combinations and weightings of a relatively small number of original test scores.

Garrett[1] has carried the work done by Hull and Limp still further, and has discovered that correlations between tests of differential aptitude increase with age, thus increasing the value of any omnibus IQ test with increasing chronological age.

Hull[2], an early worker in the field of differential

1/H. E. Garrett, "A Developmental Theory of Intelligence", American Psychologist, Vol. 1, No. 9, September, 1946, pp. 372 - 378.

2/D. L. Hull, "Variability in Amount of Different Traits Possessed by the Individual", Journal of Educational Psychology, Vol. 18, No. 2, February, 1927, pp. 97 - 106.

psychology, as early as 1925 felt that there was no single test which could be used for aptitude purposes, but that a battery of tests was essential.

In general, those in the field seem to agree that not only do individuals differ from one another, but that traits, abilities, and aptitudes of an individual differ in much the same way. These differences within individuals can be measured, though with more difficulty than can differences between individuals. Valid instruments to measure differences between individuals are readily available, but not so instruments to measure variability of abilities and traits within the individual.

On this theory, in an attempt to fill the need for such an instrument, the Differential Aptitude Tests were built.

Scientific Aptitude Tests

For many years test makers have been striving to build aptitude tests capable of differentiating the potentially successful from the potentially unsuccessful in scholastic subjects. To this end, numerous tests have been devised -- and most discarded. Seldom, indeed, has the predictive value exceeded .50, or 14% better than chance.

Although many aptitude tests have been built, surprisingly few have been planned to predict success in the field of science, and even these have been designed for the col-

lege level. One of the first attempts at scientific aptitude testing was the physics aptitude and chemistry aptitude tests of the Iowa Placement Examinations. The areas sampled by these tests were mathematics, mathematical and technical symbolism, information of chemistry or physics, and skill in solving mathematical problems verbally expressed. Correlation with first - semester college marks ran from .50 for the chemistry test to between .61 and .72 for the physics test. These tests are still in use, but the trend seems to be toward batteries measuring various factors.

Probably the test which has received the widest recognition is that test developed by Zyve[1] in 1927 -- the Stanford Scientific Aptitude Test. According to Dr. Zyve[2], he was moved to build the test because he felt that

> It would be highly instructive to determine the amount of waste resulting from such facts as that about fifty percent of graduate engineers are found fifteen years after graduation in fields of endeavour having little or nothing to do with the special training which they have received in schools of engineering.

The Zyve test is composed of exercises designed to measure eleven components which the author believed were basic to scientific aptitude. These components include:

1/D. L. Zyve, "A Test of Scientific Aptitudes"; Journal of Educational Psychology, Vol. 18, No. 8, August, 1927, pp. 525 - 546.

2/ibid, p. 525.

ability to suspend judgment, ability to detect fallacies, ability to reason logically, and ability to draw correct inductions and deductions. The test is two hours in length, and was intended for senior high school and college students. Test results appear to be quite independent of acquired knowledge. Little mathematics is required. The test is a power test and is easily administered and scored -- in fact it can be scored in five minutes. The table of norms is based on 323 college students, most of whom were Stanford University freshmen. The test results do not correlate well with college marks. Many students are able to enter engineering colleges and complete their courses successfully, though their Zyve test results are well below average. The test, however, does seem to measure something besides general scholastic aptitude, since successful students in non-scientific courses often have below-average scores.

According to Mursell[1], the validation offered was not adequately analyzed. Correlations of .95, .77, and .89 were recorded with "competent judgments on the ability of 50 research students in science."[2]

1/J. L. Mursell, Psychological Testing, Longmans, Green & Company, New York, 1950, p. 240.

2/op. cit., p. 540.

Crawford[1], and Benton and Perry[2], have made some studies of the test. Benton and Perry found correlations between test scores and college science course grades of freshmen and sophomores to be .404 ± .09 (N = 47), and of juniors and seniors .345 ± .09 (N = 43). The test correlates, .423 ± .08 (N = 46), with average physics grades, .369 ± .085 (N = 47) with average chemistry grades, and .523 ± .07 (N = 46), with average biology grades.

According to Crawford[3], the Zyve test measures not present scientific knowledge, but the aptitude for learning scientific subjects. The eleven sub-test scores are weighted, and contribution to total score varies from ratios of 2 to 7. Crawford feels that there are too few problems per sub-test, and that the validity data are not too good. The validity, as checked at Yale on 143 freshmen, was .30 with year-end grades in science or pre-engineering courses. Reliability (split-half) is .60 for the entire test, with sub-test reliabilities much lower -- though how much lower is not reported. In general, Crawford is not in favor of the

1/A. B. Crawford, "Stanford Scientific Aptitude Test", 1940 Mental Measurements Yearbook, Highland Park, New Jersey, pp. 453 - 455.

2/A. L. Benton, and J. D. Perry, "A Study of the Predictive Value of the Stanford Scientific Aptitude Test", Journal of Psychology, Vol. 10, 1940, pp. 309 - 312.

3/op. cit., p. 454.

test in its present form, and feels that it should be worked over and greatly revised.

This was the most noteworthy test in the field of scientific aptitude testing until the introduction of the Differential Aptitude Tests in 1947. The basic idea -- the determination of certain separate abilities -- is reflected in the D.A.T. Both tests are steps in a direction which may eventually prove extremely fruitful. As yet, however, neither test is as valid as would be desirable for reasonably accurate prognosis.

Previous Validation Studies of the DAT

Eight previous validation studies of the Differential Aptitude Tests are reported in the manual. These studies, according to the authors[1] are representative of the forecasting value of the tests, in that the tests were given in each case before any learning had taken place.

In the first Research Report[2] -- September, 1948 -- the authors state:

> For each school and grade listed, all correlations are given whenever data was available for fifty or more students....Correlation coefficients have been computed independently for each sex to avoid the distortions that result from combining dissimilar groups. Where the course of study appeared to be uniform for different

1/ Op. cit., p. E-11.

2/ Op. cit., p. E-11.

sections of a given grade within a city, these sections have been combined.

One consequence of this method of reporting validational data is to give many validity coefficients for each of the eight Differential Aptitude Tests.

Seven school systems have submitted data regarding the validity of the Differential Aptitude Tests for prediction in ninth grade general science. These data is summarized in Tables 1 and 2. Table 1 gives the correlation coefficients between Differential Aptitude Test scores and ninth grade general science year-end course grades for boys.[1] Table 2 gives the coefficients for girls.[2]

Table 1. Validity Coefficients between Differential Aptitude Test Scores and Year-end Course Grades in Ninth Grade General Science (Boys)

Place	N	Verb. Reas.	Num. Abil.	Abs. Reas.	Spa. Rel.	Mech. Reas.	Cler-ical	LU-I Spell.	LU-II Sent.
Worcester, Mass	59	.20	.17	.35	.26	.50	.24	-.05	.19
Cape Girardeau, Mo.	51	.66	.52	.50	.40	.42	.06	.27	.34
Columbia, Mo.	88	.80	.74	.67	.49	.50	.44	.70	.78
Independence, Mo.	142	.57	.60	.45	.34	.46	.36	.49	.54
St. Louis, Mo.	71	.67	.57	.57	.46	.37	.45	.45	.66
Mt. Vernon, N. Y.	75	.44	.43	.41	.37	.02	.17	.27	.23
Mt. Vernon, N. Y.	84	.32	.16	.28	.34	.19	-.01	.43	.43
White Plains, N. Y.	24	.55	.46	.42	.31	.37	.14	.34	.48

1/Op. cit., p. E-58.

2/Op. cit., p. E-59.

Table 2. Validity Coefficients between Differential Aptitude Test Scores and Year-end Course Grades in Ninth Grade General Science (Girls)

Place	N	Verb. Reas.	Num. Abil.	Abs. Reas.	Spa. Rel.	Mech. Reas.	Cler-ical	LU-I Spell.	LU-II Sent.
Worcester, Mass.	93	.10	.23	.16	.25	.10	.13	.10	.24
Cape Girardeau, Mo.	50	.67	.45	.50	.50	.39	.15	.36	.52
Columbia, Mo.	78	.79	.70	.70	.54	.47	.55	.65	.68
Independence, Mo.	125	.59	.63	.58	.39	.44	.50	.40	.58
Rolla, Mo.	63	.74	.72	.67	.53	.50	.52	.55	.77
St. Louis, Mo.	69	.36	.42	.20	.19	.20	.28	.27	.40
Mt. Vernon, N. Y.	91	.44	.43	.36	.39	.26	.30	.30	.38
Mt. Vernon, N. Y.	88	.45	.44	.32	.19	-.03	.22	.29	.42

The city of Worcester, Massachusetts, has utilized the Differential Aptitude Tests as an integral part of its testing program and has developed its own norms. In addition to the Worcester data included in Table 1[1/] the following were recently obtained:[2/]

1/Op. cit., p. E-58.

2/Unpublished data obtained from the office of Dr. Christensen, Director of Guidance of the Worcester public schools.

Table 3. Validity Coefficients between Differential Aptitude Test Scores and Year-end Grades in Ninth Grade General Science at Commerce High School, Worcester, Massachusetts

Sex	N	Verb. Reas.	Num. Abil.	Abst. Reas.	Space Rel.	Mech. Reas.	Clerical	LU-I Spell.	LU-II Sent.
Boys	78	.31	.33	-.02	.21	.29	.11	.12	.44
Girls	113	.55	.50	.40	.41	.41	.27	.38	.49

Undoubtedly other school systems have obtained their own norm data for their own purposes. Unfortunately, however, these data are not obtainable.

Implications of Review of Research

This thesis was predicated on the belief that there was a need for further validation of the Differential Aptitude Tests in relation to prognosis in ninth grade general science. The review of research substantiates this view.

CHAPTER III

PROCEDURES

Description of the Differential Aptitude Tests

The test battery consists of eight tests printed in seven reusable booklets. Answer sheets are provided which may be scored either by machine or by hand using a punched key. Two forms, A and B, are available. Form A was used in this study. The maximum allowable working time ranges from six minutes to thirty-five minutes, and the tests may be given singly, in toto, or any number at a time. Directions for administering and scoring the tests, individual report forms, and a manual for interpretation are also available.

According to the authors[1], the purpose of the test is

> ....to provide an integrated, scientific, and well-standardized procedure for measuring the abilities of boys and girls in grades eight through twelve for purposes of educational and vocational guidance.
>
> The tests are designed to appraise fundamental intellectual abilities, and to avoid as much as possible dependence on particular school subjects.....It will be noted that difficulty is not secured by using esoteric items or terms encountered only by those who have taken certain advanced courses. While no clear demarcation between aptitude and achievement is possible, it is believed that these tests meet the standards usually prescribed for aptitude tests as contrast-

1/Op. cit., pp. A-1, A-8.

ed with tests which attempt to appraise school learning in specific subjects.

Verbal Reasoning.-- This test contains fifty items with a score value of one point each. The score is obtained by counting the number of correct items. Maximum working time is thirty minutes. The test is designed as a power test. The authors[1] explain that the test is intended as

> ....a measure of ability to understand concepts framed in words. It is aimed at the evaluation of the student's ability to abstract or generalize and think constructively, rather than at simple fluency or vocabulary recognition. The analogies form of test item is particularly appropriate for the measurement of reasoning ability. The particular type of analogies item devised for this test is especially useful, because it provides: (1) a highly reliable item, (2) a very versatile item, and (3) a measure of reasoning that is relatively complex without being "tricky" or esoteric.

Numerical Ability.-- This test contains forty items with a score value of one point each. The score is obtained by subtracting one quarter of the number of wrong answers from the number of correct answers ($R - \frac{1}{4}W$). A Right and a Wrong answer key are provided. Maximum working time is thirty minutes. The test is designed as a power test. According to the authors[2], this test is designed

> ....to test understanding of numerical relationships and facility in handling numerical concepts. The problems are framed in the item-type form usually called "arithmetic computation" rather than in what is usually

1/Op. cit., p. A-8.

2/Op. cit., pp. A-8, A-9.

> called 'arithmetic reasoning'. This was prompted by the desire to avoid the language elements of the so-called arithmetic reasoning problem, in which reading ability may play a significant role....Many of the items, however, call for real understanding of numerical relationships; though computationally simple, they are, as problems, fully as complex as items usually framed in verbal terms.

Abstract Reasoning.-- This test consists of fifty pictorial items. Maximum score is fifty. The score is obtained by subtracting one quarter of the number of wrong answers from the number of correct answers ($R - \frac{1}{4}W$). A Right and a Wrong answer key are provided. Maximum working time is twenty-five minutes. The test is designed as a power test. The authors[1] describe the test as follows:

> The ABSTRACT REASONING test is intended as a non-verbal measure of the student's reasoning ability. The series presented in each problem requires the perception of an operating principle in the changing diagrams. In each instance, the student must discover the principle or principles, governing the change of figures and give evidence of his understanding by designating the diagram which should logically follow.....Complexity is obtained from increasing conceptual difficulty. The differences are apparent; discerning why the patterns differ is the intellectual exercise.

Space Relations.-- This test consists of forty pictorial items. Maximum score is one hundred, as more than one answer is correct for each item. The score is obtained by subtracting the number of wrong answers from the number of correct answers ($R - W$). A Right and a Wrong answer key are

1/Op. cit., p. A-9.

provided. Maximum working time is thirty minutes. The test is designed as a power test. According to the authors[1]

> The ability to visualize a constructed object from a picture of a pattern has been used frequently in tests of structural visualization. Similarly, the ability to imagine how an object would appear if rotated in various ways has been used effectively in the measurement of space perception. The item type used combines the function of these previous item types, since both factors are considered important in any useful definition of ability to think in spatial terms.
>
> A feature inherent in these items is that the test requires mental manipulation of objects in three dimensional space. Item forms which refer only to two-dimensional space are less useful, since there are relatively few occasions when perception of two-dimensional space alone is important.

Mechanical Reasoning.-- This test consists of sixty-eight pictures, with a question regarding each picture. Maximum score is sixty-eight. The score is obtained by subtracting one half the number of wrong answers from the number of correct answers ($R - \frac{1}{2}W$). A Right and a Wrong answer key are provided. Maximum working time is thirty minutes. The test is designed as a power test. The authors[2] state that

> Each item consists of a pictorially presented mechanical situation together with a simply worded question. Care has been taken to present items in terms of simple, frequently encountered mechanisms that do not resemble text-book illustrations or require special knowledge.

Clerical Speed and Accuracy.-- This test consists of two

1/Op. cit., p. A-9.

2/Op. cit., pp. A-9, A-10.

parts, each containing one hundred items. The first part is a practice exercise and is not scored. The maximum score on the second part is one hundred. The score is obtained simply by counting the number of correct responses. Maximum working time is three minutes on each part. The group being tested is not advised that the first part is for practice only. This test is the only one in the battery designed as a speed test. The authors[1] state that the test is intended to

> ....measure speed of response in a simple perceptual task. The item type desired was one which contained numbers and letters, and which required the student to go from one page to another, which involves retention over short periods of time. The item type desired possesses these characteristics.....The item thus provides a series of situations which approximate the elements involved in many clerical jobs.....There is little or no intellectual difficulty injected, since "intelligence" components are adequately measured by other tests of this battery. In this test the objective is to measure speed of perception, momentary retention, and speed of response.

Language Usage I: Spelling.-- This test consists of a list of one hundred words, some of which are spelled correctly, some incorrectly. The maximum score is one hundred. The score is obtained by subtracting the number of wrong answers from the number of correct answers (R - W). A Right and a Wrong answer key are provided. Maximum working time is ten minutes. The test is designed as a power test. The authors[2]

1/Op. cit., p. A-10.

2/Op. cit., p. A-10.

state:

> It has been known that many words are not good test items when properly spelled -- almost everyone recognizes that they are correct. Special studies of a large number of words resulted in the discovery of enough words for the test which were effective items when correctly spelled. Every item in the test is thus contributing its appropriate share of measurement; there is no 'padding'. This fact enhances the reliability of the spelling test.

Language Usage II: Sentences.-- This test consists of fifty sentences, each divided into five parts which are identified by letters. This test is in the same booklet with the spelling test, and is scheduled at the same time. The maximum score is 95 for Form A and 94 for Form B. The score is obtained by subtracting the number of wrong answers from the number of correct answers (R - W). A Right and a Wrong answer key are provided. Maximum working time is twenty-five minutes. The test is designed as a power test. The authors[1] state:

> The SENTENCES section of the LANGUAGE USAGE test is intended to measure the student's ability to distinguish between good and bad grammar, punctuation and word usage....The student must inspect each part of each sentence and judge whether or not it is correct. Since there may be any number of errors in each sentence, the student cannot guess by elimination of parts of sentences. He must react separately to each part of each sentence.

Description of the Criteria

Description of the Group Tested.-- The two groups used in

1/Op. cit., p. A-5.

this study include the entire ninth grade class of the Bigelow Junior High School, Newton, Massachusetts, and all ninth grade students in the city of Worcester, Massachusetts, taking general science. Data was obtained on forty boys and thirty-five girls in Newton, and twenty-two boys and forty-nine girls in Worcester; a total of sixty-two boys and eighty-four girls, or one hundred and forty-six cases in all.

The general science course in Newton is a required course, while in Worcester it is an elective course.

According to the authors[1], considerable sex differences have been observed in the results of some of the tests, so it was decided to brēak down the results of the study by sexes. In addition, it was felt that since the groups used were not random samples, and since they differed both in central tendency and variability of the measure, it would also be advisable to provide a break-down of the results according to the two groups. Results therefore were determined for Newton boys, girls, and boys and girls; Worcester boys, girls, and boys and girls; and the total results for boys, girls, and boys and girls.

The IQ's of the two groups were close to, though slightly above average as measured by the Otis DM. Data regarding central tendency and variability of all the above groups on

1/Op. cit., p. A-5

the Otis test are shown in Table 4.

Table 4. Variability and Central Tendency of Otis DM IQ's for All Groups in the Study

Place	Group	N	Mean	S.D.	σ_m
Newton	Boys	33	106.73	12.12	2.14
	Girls	27	109.78	10.47	2.06
	Boys and Girls	60	107.75	11.97	1.55
Worcester	Boys	22	107.05	8.41	1.83
	Girls	49	100.58	10.34	1.49
	Boys and Girls	71	102.60	10.67	1.27
Newton and Worcester	Boys	55	106.51	11.31	1.54
	Girls	76	103.84	11.29	1.29
	Boys and Girls	131	104.96	11.38	.99

It will be noted that the Worcester group tends to be somewhat more homogeneous in IQ than does the Newton group. This held true for about all of the Differential Aptitude Test results also. The difference in homogeneity is probably accounted for by the fact that in Newton the sample included the total population of the ninth grade of the school, both college and non-college courses, while the Worcester sample was almost exclusively non-college preparatory.

Description of the marking systems.-- To render both sets of marks comparable, the Newton grades A, B, C, D and E were assigned values of 4, 3, 2, 1 and 0 respectively, and these, together with the percentages used by Worcester, were reduced

to standard scores, each pupil being assigned two standard scores for each grade, one based upon the local distribution for his sex, and one upon the distribution of both sexes together. The formula given by Lindquist[1] was used: $T = \frac{10\,(X-M)}{S.D.} + 50$, where T = Standard Score, M = mean of the distribution of raw scores, S.D. = the standard deviation of the raw score distribution, and X is a particular raw score in the distribution. Each course mark was that given at the end of the year as a final cumulative grade.

Description of Read General Science Test.[2] -- Recognizing the subjectivity of teachers' marks, it was deemed advisable to use some more objective criterion. To this end, the Read Test was selected as being the best obtainable objective measure of general science achievement. Of the two forms available, AM and BM, the form AM was used.

The test consists of seventy-five multiple choice questions, with a selection of five responses for each question. The test utilized a reusable test booklet with a separate answer sheet, which may be either hand or machine scored. The score is obtained simply by counting the number of correct responses, using the punched key which is provided.

1/E. F. Lindquist, A First Course in Statistics, Houghton Mifflin Company, Cambridge, Massachusetts, 1942, p. 149.

2/J. G. Read, Read General Science Test, World Book Company, Yonkers-on-Hudson, New York, 1951.

Scores are convertible to standard scores and percentiles. The standard scores, however, have a mean of 102 and a S.D. of 12.5, so chosen because they represented the median and S.D. of the distribution of Terman-McNemar IQ's of the students in the norm sample. The maximum working time is forty minutes and the maximum possible score is 75.

Administration of Tests.-- The Differential Aptitude Tests were administered in the Fall of 1950 shortly after the beginning of the school year. The tests were administered by the individual schools. Though this is a possible limitation, the administrators involved claim to have followed precisely the standardized procedures set forth in the Manual.[1/] The Read General Science Test was administered by the individual science teachers in June, 1951, following the standardized procedures set forth in the accompanying manual.

Description of Statistical procedures Used.-- The mean, standard deviation, and standard error of the mean were found for the raw scores of each test in the battery, for the raw score on the total battery, for each course mark, and for the raw scores on the Read General Science Achievement Test.

Correlations.-- Intercorrelation coefficients for the Differential Aptitude Test battery, correlation coefficients

1/Op. cit., pp. B-1 -- B-8.

between each test variable and teachers' final marks in English and general science, and correlation coefficients between each test variable and the Read General Science Test results were calculated, using the Pearson Product Moment formula suggested by Lindquist[1]:

$$r_{xy} = \frac{\frac{\Sigma XY}{N} - M_x M_y}{\sigma_x \sigma_y}$$

Multiple correlation coefficients were also computed using several pairs of the Differential Aptitude Test scores with the hope of finding a combination with a high R in relation to Read General Science Test raw scores. The formula used was suggested by Peatman[2]:

$$R_{c \cdot xy} = \sqrt{\frac{r_{cx}^2 + r_{cy}^2 - 2 r_{cx}\, r_{cy}\, r_{xy}}{1 - r_{xy}^2}}$$

Multiple correlation coefficients were calculated using as the primary consideration low intercorrelations, with relatively high correlations with the Read General Science Test scores as the secondary consideration.

Summary of Procedures

In order to determine statistically the prognostic value of the Differential Aptitude Tests in predicting suc-

1/Op. cit., pp. 167 - 168.

2/F. G. Peatman, Descriptive and Sampling Statistics, Harper & Brothers, New York, 1947, p. 483.

cess or failure in English and general science in the ninth grade of the Worcester and Newton junior high schools, it was considered advisable to include:

1. A statistical description of the central tendency and variability of the Otis IQ's of the groups involved.

2. A statistical description of the central tendency, variability and intercorrelations of the Differential Aptitude Test scores of the groups involved.

3. A statistical description of the central tendency, variability, and intercorrelations of the criteria, i.e., teachers' final marks and Read General Science Test raw scores, for each group involved.

4. Correlation coefficients between each Differential Aptitude Test score and the criteria for each group.

5. Finally, multiple correlations, using combinations of selected pairs of the Differential Aptitude Tests and Read General Science Test raw scores for the combined groups only.

Chapter IV

RESULTS OF STUDY

Statistical Description of Raw Scores

Tables 5, 6, and 7 show the central tendency and variability of each Differential Aptitude Test and of the total battery, for Newton, Worcester, and combined Newton and Worcester. Also shown for the sake of comparison are means, S.D.'s, and N's given in the table of norms of the Manual[1/] for grade nine.

In using the raw scores in the following description, the procedure followed by the authors of the Differential Aptitude Tests has been used, i. e., all negative scores have been considered to be zero.

It will be noted that the tables indicate quite clearly that the Newton group is slightly more heterogeneous than the norm groups, while the Worcester group, particularly the girls, are considerably more homogeneous than are the norm groups. With the combined groups, the boys approximate the mean and S.D. of the norm group very closely, except for Space Relations and Mechanical Reasoning, in which the S.D.'s are close to those of the norm group, though

1/Op. cit., p. D-8

Table 5. Central Tendency and Variability of Differential Aptitude Test Raw Scores in the Ninth Grade of the Bigelow Junior High School, Newton, Massachusetts, Compared with the Norm Group as Reported in the Manual 1/

Group	Newton Group					Norm Group		
	Test	N	Mean	σm	S.D.	N	Mean	S.D.
Boys	Verb. Reas.	40	20.62	1.42	8.88	1816	19.5	8.6
	Num. Abil.	40	19.50	1.59	9.88	1828	17.2	8.1
	Abs. Reas	40	27.38	1.48	9.18	1834	25.0	10.5
	Space Rel.	40	39.25	3.74	23.31	1848	40.1	23.8
	Mech. Reas.	40	31.00	2.14	13.36	1835	35.8	12.6
	Cler. S & A	40	49.12	2.00	12.14	1748	46.8	10.5
	LU-I Spell.	40	34.50	4.13	25.76	1817	33.1	24.1
	LU-II Sent.	40	26.62	2.14	13.39	1815	25.6	14.8
	Total DAT	40	247.50	14.17	88.26	----	----	----
Girls	Verb. Reas.	35	21.86	1.66	9.65	1927	19.4	8.8
	Num. Abil.	35	19.71	1.40	8.14	1953	16.6	7.7
	Abs. Reas.	35	28.86	1.67	9.72	1936	23.4	11.0
	Space Rel.	35	40.57	4.13	24.02	1964	32.9	20.3
	Mech. Reas	35	22.14	1.58	9.22	1944	21.4	10.6
	Cler. S & A	35	56.86	1.54	8.98	1855	54.2	11.8
	LU-I Spell.	35	46.00	3.95	22.95	1969	46.1	24.2
	LU-II Sent.	35	36.57	2.71	15.74	1969	33.3	15.1
	Total DAT	35	271.07	14.65	85.32	----	----	----
Boys and Girls	Verb. Reas.	75	21.20	1.04	8.95	----	----	----
	Num. Abil.	75	19.60	1.05	9.10	----	----	----
	Abs. Reas.	75	28.05	1.10	9.50	----	----	----
	Space Rel.	75	39.54	2.74	23.60	----	----	----
	Mech. Reas.	75	26.87	1.29	11.15	----	----	----
	Cler. S.& A	75	52.73	1.32	11.40	----	----	----
	LU-I Spell.	75	40.36	2.86	24.75	----	----	----
	LU-II Sent.	75	31.27	1.80	15.60	----	----	----
	Total DAT	75	258.50	10.22	88.40	----	----	----

1/ Op. cit., p. D-8

Table 6. Central Tendency and Variability of Differential Aptitude Test Raw Scores in the Ninth Grade of Several Worcester, Massachusetts, Junior High Schools Compared with the Norm Group as Reported in the Manual 1/

Group	Worcester Group					Norm Group		
	Test	N	Mean	SM	S.D.	N	Mean	S.D.
Boys	Verb.Reas.	22	18.59	1.52	6.98	1816	19.5	8.6
	Num. Abil.	22	13.36	1.55	7.10	1828	17.2	8.1
	Abs. Reas.	22	22.00	2.28	10.44	1834	25.0	10.5
	Space Rel.	22	32.22	4.60	21.07	1845	40.1	23.8
	Mech. Reas.	22	31.32	2.00	9.20	1835	35.8	12.6
	Cler.S & A	22	42.91	2.46	11.24	1748	46.8	10.5
	LU-I Spell	22	34.50	3.65	16.70	1817	33.1	24.1
	LU-II Sent	22	24.27	1.85	8.49	1815	25.6	14.8
	Total DAT	22	218.68	11.58	53.14	----	----	----
Girls	Verb.Reas.	49	14.14	0.85	5.89	1927	19.4	8.8
	Num.Abil.	49	12.71	0.88	6.06	1953	16.6	7.7
	Abs. Reas.	49	18.84	1.61	11.19	1936	23.4	11.0
	Space Rel.	49	28.02	2.46	17.02	1964	32.9	20.3
	Mech. Reas.	49	18.94	1.29	8.91	1944	21.4	10.6
	Cler. S & A	49	49.45	1.88	13.02	1855	54.2	11.8
	LU-I Spell.	49	30.78	2.79	19.31	1969	46.1	24.2
	LU-II Sent.	49	23.64	1.89	13.12	1969	33.3	15.1
	Total DAT	49	195.40	8.06	55.76	----	----	----
Boys and Girls	Verb.Reas.	71	15.52	0.79	6.58	----	----	----
	Num.Abil.	71	12.92	0.77	6.41	----	----	----
	Abs. Reas.	71	19.82	1.32	11.06	----	----	----
	Space Rel.	71	29.32	2.21	18.56	----	----	----
	Mech. Reas.	71	22.28	1.46	12.22	----	----	----
	Cler. S & A	71	47.42	1.54	12.86	----	----	----
	LU-I Spell.	71	32.50	2.26	18.92	----	----	----
	LU-II Sent.	71	23.83	1.42	11.88	----	----	----
	Total DAT	71	202.62	6.70	56.02	----	----	----

1/ Op. cit., p. D-8

Table 7. Central Tendency and Variability of Differential Aptitude Test Raw Scores in the Ninth Grades of the Bigelow Junior High School, Newton, Massachusetts and Several Worcester, Massachusetts, Junior High Schools Compared with the Norm Group as Reported in the Manual 1/

Group	Combined Newton and Worcester Groups					Norm Group		
	Test	N	Mean	σm	S.D.	N	Mean	S.D.
Boys	Verb Reas.	62	19.90	1.06	8.30	1816	19.5	8.6
	Num. Abil.	62	17.32	1.22	9.46	1828	17.2	8.1
	Abs. Reas.	62	25.47	1.28	9.98	1834	25.0	10.5
	Space Rel.	62	36.76	2.96	23.16	1848	40.1	23.8
	Mech. Reas.	62	30.71	1.55	12.08	1835	35.8	12.6
	Cler. S & A	62	46.92	1.56	12.20	1748	46.8	10.5
	LU-I Spell.	62	34.50	2.94	22.96	1817	33.1	24.1
	LU-II Sent.	62	25.79	1.53	11.94	1815	25.6	14.8
	Total DAT	62	237.60	10.11	78.82	----	----	----
Girls	Verb. Reas.	84	17.24	0.90	8.20	1927	19.4	8.8
	Num. Abil.	84	15.63	0.85	7.80	1953	16.6	7.7
	Abs. Reas.	84	23.01	1.28	11.70	1936	23.4	11.0
	Space Rel.	84	33.13	2.32	21.31	1964	32.9	20.3
	Mech. Reas.	84	19.68	1.06	9.68	1944	21.4	10.6
	Cler. S & A	84	52.54	1.32	12.08	1855	54.2	11.8
	LU-I Spell.	84	37.60	2.44	22.38	1969	46.1	24.2
	LU-II Sent.	84	29.02	1.72	15.70	1969	33.3	15.1
	Total DAT	84	226.92	8.64	79.06	----	----	----
Boys and Girls	Verb. Reas.	146	18.92	0.65	7.87	----	-----	----
	Num. Abil.	146	16.35	0.71	8.59	----	----	----
	Abs Reas.	146	24.06	0.92	11.07	----	----	----
	Space Rel.	146	34.67	1.84	22.22	----	-----	----
	Mech Reas.	146	24.36	1.00	12.07	----	----	----
	Cler. S & A	146	50.15	1.03	12.44	----	----	----
	LU-I Spell	146	36.59	1.86	22.44	----	----	----
	LU-II Sent.	146	27.65	1.18	14.26	----	----	----
	Total DAT	146	231.32	6.55	79.14	----	----	----

1/ Op. cit., p. D-8

the means are somewhat lower. With the girls, on the other hand, while the S.D.'s tended to approximate those of the norm group, the means of the two Language Usage tests were considerably below those of the norm group. This is due entirely to the low scores obtained by the Worcester girls. This is probably accounted for by a number of factors, one of the largest of which may be the fact that the Worcester girls were all taking non-college preparatory courses, and the tendency to use the commercial and home economics programs as "dumping grounds" for the scholastically undesirable is well known. This is further borne out by the relatively low IQ's of this group (See Table 4, page 28).

Intercorrelation of the Differential Aptitude Tests

The intercorrelation coefficients for the Newton Groups are found in Tables 8, 9, and 10.

Table 8. Intercorrelation Coefficients between the Differential Aptitude Test Raw Score Results for 40 Ninth Grade Boys, Bigelow Junior High School, Newton, Massachusetts

Test	Verb. Reas.	Num. Abil.	Abs. Reas.	Space Rel.	Mech Reas.	Cler. S&A	LU-I Spell	LU-II Sent.
Total DAT	.70	.74	.88	.82	.54	.50	.78	.73
LU-II Sent.	.51	.56	.76	.46	.22	.11	.63	
LU-I Spell.	.56	.64	.60	.62	.08	.19		
Cler. S & A	.22	.46	.73	.30	.06			
Mech. Reas.	.32	.17	.46	.45				
Space Rel.	.62	.53	.62					
Abs. Reas.	.60	.71						
Num. Abil.	.57							

It is interesting to note the wide range of the coefficients within each table. For example, the coeffients for the boys, as shown in Table 8, range from a low of .06 to a high of .88. As might be expected, the correlations of the sub-tests with the total score are much higher than those between sub-tests. Mechanical Reasoning is of special interest due to its wide range of .06 to .54.

Table 9. Intercorrelation Coefficients between the Differential Aptitude Test Raw Score Results for 35 Ninth Grade Girls, Bigelow Junior High School, Newton, Massachusetts.

Test	Verb. Reas.	Num. Abil.	Abs. Reas.	Space Rel.	Mech. Reas.	Cler. S&A	LU-I Spell	LU-II Sent.
Total DAT	.76	.70	.67	.78	.66	.62	.74	.88
LU-II Sent.	.78	.60	.53	.34	.58	.48	.65	
LU-I Spell.	.68	.46	.32	.32	.34	.43		
Cler. S & A.	.56	.30	.29	.35	.18			
Mech Reas.	.46	.31	.37	.60				
Space Rel.	.56	.44	.55					
Abs. Reas.	.53	.47						
Num. Abil.	.50							

Comparison of Tables 8 and 9 shows considerably less variability in the intercorrelation coefficients for the girls than for the boys. According to these results, the individual girls appeared to be more homogeneous in the qualities measured by the sub-tests than did the boys. While there appears to be little relationship between score on the Mechanical Reasoning test and most of the other tests in the

battery among the boys, the relationship was fairly high among the girls.

Table 9 shows that the correlations of the total score with the sub-tests are in most cases considerably higher than between sub-tests.

Table 10. Intercorrelation Coefficients between the Differential Aptitude Test Raw Score Results for 75 Ninth Grade Boys and Girls, Bigelow Junior High School, Newton, Massachusetts.

Test	Verb. Reas.	Num. Abil.	Abs. Reas.	Space Rel.	Mech. Reas.	Cler. S&A	LU-I Spell.	LU-II Sent.
Total DAT	.81	.71	.78	.83	.46	.55	.70	.78
LU-II Sent.	.65	.52	.62	.54	.22	.34	.62	
LU-I Spell.	.59	.52	.43	.45	.03	.25		
Cler. S & A	.37	.38	.55	.37	.12			
Mech. Reas.	.37	.22	.40	.54				
Space Rel.	.64	.52	.63					
Abs. Reas.	.59	.60						
Num. Abil.	.69							

Table 10 shows the intercorrelation coefficients for the combined groups of boys and girls. As might be expected, most of the coefficients fall between those found for the separate groups. There are a few exceptions, however, such as the coefficients between total score and Verbal Reasoning or Space Relations. There is also a noticeable tendency for the variability of the coefficients to decrease markedly.

The intercorrelation coefficients for the Worcester groups are found in Tables 11, 12, and 13.

Table 11. Intercorrelation Coefficients between the Differential Aptitude Test Raw Score Results for the 22 Ninth Grade Boys of Several Worcester, Massachusetts, Junior High Schools.

Test	Verb. Reas.	Num. Abil.	Abs. Reas.	Space Rel.	Mech. Reas.	Cler. S& A	LU-I Spell.	LU-II Sent.
Total DAT	.34	.43	.49	.63	.45	.28	.08	.71
LU-II Sent.	.59	.14	.40	.41	.57	.18	.16	
LU-I Spell.	-.08	.06	-.13	-.23	-.34	-.20		
Cler. S & A	.16	.20	.22	.31	.18			
Mech. Reas.	.47	.24	.39	.21				
Space Rel.	.10	.31	.28					
Abs. Reas.	.13	.22						
Num. Abil.	.08							

The difference between the results shown in Table 8 and those shown in Table 11 are very interesting. Negative correlations as high as -.34 are found in Table 11, while all correlations in Table 8 are positive. Almost all of the correlations in Table 11 are very considerably lower than those shown in Table 8. Apparently there is much less relationship among the Worcester boys in those qualities measured by the various sub-tests than among the Newton boys.

Table 12. Intercorrelation Coefficients between the Differential Aptitude Test Raw Score Results for 49 Ninth Grade Girls of Several Worcester, Massachusetts, Junior High Schools.

Test	Verb. Reas.	Num. Abil.	Abs. Reas.	Space Rel.	Mech. Reas.	Cler. S & A	LU-I Spell.	LU-II Sent.
Total DAT	.49	.34	.63	.55	.25	.13	.58	.75
LU-II Sent.	.53	.37	.39	.19	.06	-.01	.60	
LU-I Spell.	.33	.17	.17	-.08	.03	-.16		
Cler. S & A	-.16	-.16	.01	.21	-.31			
Mech. Reas.	.14	-.06	.34	.16				
Space Rel.	.23	.20	.49					
Abs. Reas.	.25	.18						
Num. Abil.	.12							

The differences between the results shown in Table 9 and those shown in Table 12 illustrate considerable difference between the Newton and Worcester girls. The correlations in Table 12 are very low, and there appears to be very little relationship between the sub-tests. It may again be noted, however, that the highest correlations are between total score and the sub-test scores, and not between sub-tests in most cases.

Table 13. Intercorrelation Coefficients Between The Differential Aptitude Test Raw Score Results for 71 Ninth Grade Boys and Girls of Several Worcester, Massachusetts, Junior High Schools

Test	Verb. Reas.	Num. Abil.	Abs. Reas.	Space Rel.	Mech. Reas.	Cler. S & A	LU-I Spel.	LU-II Sent.
Total DAT	.46	.36	.60	.58	.46	.12	.34	.72
LU-II Sent.	.51	.30	.39	.24	.22	.03	.44	
LU-I Spell.	.16	.07	.05	-.16	.01	-.35		
Cler. S & A	-.22	-.06	.03	.15	.08			
Mech. Reas.	.42	.13	.40	.24				
Space Rel.	.20	.24	.42					
Abs. Reas.	.24	.18						
Num. Abil.	.11							

With few exceptions, the results in Table 13 show most of the correlations falling between those for the boys and those for the girls, as was to be expected. The coefficient between Clerical Speed and Accuracy and Language Usage-I: Spelling, shows a definite negative relationship. This was interesting, as the combined Newton groups showed almost as high a positive relationship.

It will be noted that in some cases a wide range of correlation coefficinets are found among like correlations in the various groups. This is due, at least to some extent, to the small size of N in some groups. A summary table, Table XIII, showing the median r's will be found on page in the appendix.

Because of the small N's, the use of the standard error

Table 14. Intercorrelation coefficients between the Differential Aptitude Test Raw Score Results for 62 Ninth Grade Boys of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Test	Verb. Reas.		Num. Abil.		Abst. Reas.		Space Rel.		Mech. Reas.		Cler. S & A		LU-I Spell.		LU-II Sent.	
	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr
Total DAT	.62	.08	.69	.07	.74	.06	.75	.06	.50	.09	.45	.10	.64	.08	.72	.06
LU-II Sent.	.53	.09	.47	.10	.63	.08	.44	.10	.36	.11	.14	.12	.54	.09		
LU-I Spell.	.42	.10	.50	.09	.36	.11	.40	.10	.04	.12	.14	.12				
Cler. S & A.	.14	.12	.43	.10	.57	.09	.32	.11	.13	.12						
Mech. Reas.	.43	.10	.23	.12	.50	.10	.07	.12								
Space Rel.	.47	.10	.48	.10	.50	.10										
Abs. Reas.	.46	.10	.58	.09												
Num. Abil.	.46	.10														

Table 15. Intercorrelation Coefficients between the Differential Aptitude Test Raw Score Results for 84 Ninth Grade Girls of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Test	Verb. Reas.		Num. Abil.		Abst. Reas.		Space Rel.		Mech. Reas.		Cler. S & A		LU-I Spell.		LU-II Sent.	
	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr
Total DAT	.87	.03	.64	.06	.70	.06	.73	.05	.63	.07	.41	.09	.64	.06	.85	.03
LU-II Sent.	.78	.04	.58	.08	.54	.08	.49	.08	.44	.09	.27	.10	.63	.07		
LU-I Spell.	.59	.07	.37	.09	.30	.10	.20	.10	.27	.10	.06	.11				
Cler. S & A	.33	.10	.15	.11	.20	.10	.29	.10	.19	.10						
Mech. Reas.	.50	.08	.29	.10	.48	.08	.49	.08								
Space Rel.	.56	.08	.43	.09	.57	.08										
Abs. Reas.	.54	.08	.43	.09												
Num. Abil.	.68	.06														

Table 16. Intercorrelation Coefficients Between the Differential Aptitude Test Raw Score Results for 146 Ninth Grade Boys and Girls of the Bigelow Junior High School, Newton, Massachusetts and Several Worcester, Massachusetts, Junior High Schools

Test	Verb. Reas.		Num. Abil.		Abst. Reas.		Space Rel.		Mech. Reas.		Cler. S & A		LU-I Spell.		LU-II Sent.	
	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr
Total DAT	.77	.03	.66	.05	.60	.05	.74	.04	.54	.06	.41	.07	.60	.05	.79	.03
LU-II Sent.	.56	.06	.51	.06	.55	.06	.45	.07	.30	.07	.24	.07	.58	.05		
LU-I Spell.	.39	.07	.47	.06	.28	.07	.26	.07	.08	.08	.01	.08				
Cler. S & A	.06	.08	.25	.07	.31	.07	.28	.07	.04	.08						
Mech. Reas.	.37	.07	.28	.07	.48	.06	.44	.07								
Space Rel.	.44	.07	.45	.07	.54	.06										
Abs. Reas.	.39	.07	.49	.06												
Num. Abil.	.48	.06														

of r was not valid, except for the combined cases of Worcester and Newton. According to Lindquist,[1] the standard error of r should be used only with N's over 60.

Tables 14, 15 and 16 show the intercorrelation coefficients and the standard errors of the coefficients for the combined Worcester and Newton groups.

It is apparent from Tables 14, 15 and 16 that there are definite relationships between the various sub-tests. Since this is the case, the factors measured by the various sub-tests must overlap somewhat, with the possible exception of Clerical Speed and Accuracy and Mechanical Reasoning. As might be expected, the range of the coefficients is considerably less than for the smaller groups.

Statistical Description of the Criteria

Tables 17, 18 and 19 show the central tendency and variability of each of the criteria, i.e., Read General Science Test raw scores, and teachers' marks in General Science and English.

The tables indicate quite clearly that the Newton Group has achieved somewhat more in General Science than has the Worcester group, as might be expected from the Otis DM IQ results.

1/ E. F. Lindquist, A First Course in Statistics, Houghton Mifflin Company, Cambridge, Massachusetts 1942, p. 192.

Table 17. Central Tendency and Variability of Read General Science Test Raw Scores and Teachers' Marks in General Science and English (Reduced to Standard Scores) in the Ninth Grade of the Bigelow Junior High School, Newton, Massachusetts

Group	Criterion	N	Mean	σ_m	S.D.
Boys	Read General Science Test	40	43.75	1.70	10.58
	General Science Mark	40	50.00	1.60	10.00
	English Mark	40	50.00	1.60	10.00
Girls	Read General Science Test	35	37.86	1.90	11.05
	General Science Mark	35	50.00	1.72	10.00
	English Mark	35	50.00	1.72	10.00
Boys and Girls	Read General Science Test	75	41.00	1.30	11.20
	General Science Mark	75	50.00	1.16	10.00
	English Mark	75	50.00	1.16	10.00

Comparison of Tables 17, 18 and 19 shows a definite sex difference in the mean of the Read General Science Test raw score. An approximate six-point differential in favor of the boys is found in each case, though the Newton girls are somewhat more heterogeneous than the boys.

Table 18. Central Tendency and Variability of Read General Science Test Raw Scores and Teachers' Final Marks in General Science and English (Reduced to Standard Scores) in the Ninth Grades of Several Worcester, Massachusetts, Junior High Schools.

Group	Criterion	N	Mean	σ_m	S.D.
Boys	Read General Science Test	22	40.50	2.12	9.74
	General Science Mark	22	50.00	2.18	10.00
	English Mark	22	50.00	2.18	10.00

(concluded on next page)

Table 18. (concluded)

Group	Criterion	N	Mean	σ_m	S.D.
Girls	Read General Science Test	49	33.82	1.19	8.23
	General Science Mark	49	50.00	1.44	10.00
	English Mark	49	50.00	1.44	10.00
Boys and Girls	Read General Science Test	71	35.88	1.11	9.26
	General Science Mark	71	50.00	1.20	10.00
	English Mark	71	50.00	1.20	10.00

Table 19. Central Tendency and Variability of Read General Science Test Raw Scores and Teachers' Final Marks in General Science and English (Reduced to Standard Scores) in the Ninth Grades of Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Group	Criterion	N	Mean	σ_m	S.D.
Boys	Read General Science Test	62	42.60	1.33	10.40
	General Science Mark	62	50.00	1.28	10.00
	English Mark	62	50.00	1.28	10.00
Girls	Read General Science Test	84	35.50	1.15	10.45
	General Science Mark	84	50.00	1.10	10.00
	English Mark	84	50.00	1.10	10.00
Girls and Boys	Read General Science Test	146	38.52	0.88	10.61
	General Science Mark	146	50.00	0.83	10.00
	English Mark	146	50.00	0.83	10.00

Correlation of the Differential Aptitude Test Scores With Read General Science Test Scores

Table 20 gives the r's for each group tested between the raw scores of each of the Differential Aptitude Tests and the raw scores of the Read General Science Test. Stan-

Table 20. Correlation Coefficients between Differential Aptitude Test Raw Scores and Read General Science Test Raw Scores, and between Otis DM IQ's and Read General Science Test Raw Scores for All Groups Used in This Study

Group	N	Otis DM IQ		Verb. Reas.		Num. Abil.		Abst. Reas.		Space Rel.		Mech. Reas.		Cler. S & A		LU-I Spell.		LU-II Sent.		Total DAT	
		r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr
Newton																					
Boys	40	.53		.58		.48		.76		.48		.33		.28		.54		.62		.71	
Girls	35	.38		.74		.53		.37		.52		.46		.32		.46		.60		.70	
Boys and Girls	75	.45		.64		.48		.53		.59		.52		.18		.35		.46		.63	
Worcester																					
Boys	22	.72		.48		.03		.27		.29		.29		-.01		.18		.58		.47	
Girls	49	.73		.50		.21		.38		.30		-.07		-.32		.53		.68		.56	
Boys and Girls	71	.71		.55		.14		.37		.31		.35		-.28		.30		.61		.55	
Newton and Worcester																					
Boys	62	.62	.08	.56	.09	.38	.11	.59	.09	.43	.10	.50	.09	.21	.12	.44	.10	.60	.08	.64	.08
Girls	84	.83	.03	.69	.06	.40	.09	.38	.09	.44	.09	.40	.09	.01	.11	.41	.09	.61	.07	.60	.07
Boys and Girls	146	.72	.04	.45	.07	.41	.07	.48	.06	.45	.07	.54	.06	.02	.08	.35	.07	.55	.06	.63	.05

dard errors of r are shown for the combined results. Included also in the table are the r's between the Otis DM IQ's and the raw scores of the Read General Science Test.

According to Table 20, the best single predictor for Newton is the Total DAT, closely followed by Verbal Reasoning. The best predictor for Worcester is the Otis DM IQ, closely followed by Language Usage-II: Sentences. For the combined groups, the best predictor is the Otis DM IQ, with the Total DAT a close second.

It will be noted that only in the case of the Newton Group do any of the Differential Aptitude Test scores show higher correlation with the criterion than does the Otis DM IQ. Further, those tests of the battery, i.e., Verbal Reasoning, Abstract Reasoning, and both Language Usage tests, which tend to measure those dimensions of behavior measured by a general intelligence test also tend to correlate highly with the criterion. As might be expected, the scores on the total battery tended to correlate quite well with the criterion.

Multiple Correlations

According to Tables 14 and 20, it would seem that the pairs of Differential Aptitude Tests which would tend to correlate most highly with the criterion for the total group of boys would be: Mechanical Reasoning and Language Usage-I:

Spelling, Mechanical Reasoning and Space Relations, Clerical Speed and Accuracy and Language Usage-II: Sentences, and Mechanical Reasoning and Total DAT score, since they have relatively low intercorrelations and reasonably high correlations with the criterion.

Table 21. Intercorrelation Coefficients of the Differential Aptitude Tests and Multiple Correlations of Selected Pairs of Differential Aptitude Tests with Read General Science Achievement Test Raw Scores for 62 Ninth Grade Boys of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Combinations of the DAT	Inter-correlations	r with Criterion	R with Criterion	Gain
Mech. Reas. LU-I: Spell.	.04	.50 .44	.65	.15
Mech. Reas. Space Rel.	.07	.50 .43	.63	.13
Clerical LU-II: Sent.	.14	.21 .60	.61	.01
Mech. Reas. Total DAT	.50	.50 .64	.67	.03

Inspection of Table 21 shows that the best multiple predictors are Mechanical Reasoning and Total DAT. However, almost as good are Mechanical Reasoning and Language Usage-I: Spelling, with a difference of only .02.

According to Tables 15 and 20, the pairs of differential Aptitude Test scores which would tend to correlate most highly with the criterion for the total group of girls would be: Verbal Reasoning and Language Usage-II: Sentences, Verbal Reasoning and Total DAT, Language Usage-II: Sentences and Total DAT, and Verbal Reasoning and Mechanical Reasoning.

Table 22 shows the multiple correlation coefficients found for the above pairs.

Table 22. Intercorrelation Coefficients of the Differential Aptitude Tests and Multiple Correlations of Selected Pairs of Differential Aptitude Tests with Read General Science Test Raw Scores for 84 Ninth Grade Girls of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Combinations of the DAT	Inter-correlations	r with Criterion	R with Criterion	Gain
Verb. Reas. LU-II: Sent.	.78	.69 .61	.70	.01
Verb. Reas. Total DAT	.87	.69 .60	.71	.02
LU-II: Sent. Total DAT	.85	.61 .60	.62	.01
Verb. Reas. Mech. Reas.	.50	.69 .40	.69	$<.01$

It is apparent from Table 22, that there is very little increase in predictive accuracy by the use of multiple correlations. Maximum gain is only .02, hardly enough to make use of multiple predictors worthwhile.

According to Tables 16 and 20, the pairs of Differential Aptitude Test scores which would tend to correlate most highly with the criterion for the total group of boys and girls would be: Mechanical Reasoning and Language Usage-I: Spelling, Mechanical Reasoning and Language Usage-II: Sentences, Mechanical Reasoning and Total DAT, and Abstract Reasoning and Language Usage-I: Spelling.

Table 23 shows the multiple correlation coefficients found for the above pairs.

Table 23. Intercorrelation Coefficients of the Differential Aptitude Tests and Multiple Correlations of Selected Pairs of Differential Aptitude Tests with Read General Science Test Raw Scores for 146 Ninth Grade Boys and Girls of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Combinations of the DAT	Inter-correlations	r with Criterion	R with Criterion	Gain
Mech. Reas. LU-I: Spell.	.08	.54 .35	.62	.08
Mech. Reas. LU-II: Sent.	.30	.54 .55	.67	.12
Mech. Reas. Total DAT	.54	.54 .63	.67	.04
Abst. Reas. LU-I: Spell.	.28	.48 .35	.52	.04

According to Table 23, the best multiple predictors are Mechanical Reasoning and Language Usage-II: Sentences, and Mechanical Reasoning and Total DAT.

Correlation of the Differential Aptitude Test Raw Scores With Teachers' Final Marks

Tables 24 and 25 give the r's for each group tested between the raw scores of each of the Differential Aptitude Tests and the teachers' final marks (reduced to Standard Scores). Included also in the tables are the r's between the Otis DM IQ's and the teachers' final marks.

It is interesting to note that on the whole the Otis DM IQ shows about as high correlation with teachers' marks as do any of the Differential Aptitude Tests.

There are undoubtedly many factors other than achievement that enter into most teachers' marks. With this in mind, a comparison of Tables 20 and 24 shows some significant differences. For example, for the total group of girls, the Otis DM IQ correlates .84 with Read General Science Test raw scores, but only .30 with teachers' marks in general science. Also, for the total group of girls, the Total DAT correlates .60 with Read General Science Test raw scores, but only .33 with teachers' marks in general science. Almost without exception correlations are markedly higher between predictors and Read General Science

Table 24. Correlation Coefficients between Differential Aptitude Test Raw Scores and Teachers' Final Marks in General Science (Reduced to Standard Scores), and between Otis DM IQ's and Teachers' Final Marks in General Science (Reduced to Standard Scores) for All Groups Used in This Study

Group	N	Otis DM IQ		Verb. Reas.		Num. Abil.		Abst. Reas.		Space Rel.		Mech. Reas.		Cler. S & A		LU-I Spell.		LU-II Sent.		Total DAT	
		r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr
Newton																					
Boys	40	.63		.59		.63		.54		.60		.28		.46		.53		.58		.69	
Girls	35	.58		.79		.64		.41		.33		.34		.54		.49		.55		.67	
Boys and Girls	75	.64		.68		.62		.46		.51		.36		.35		.42		.46		.64	
Worcester																					
Boys	22	.34		.13		.42		.47		.24		.18		.50		.09		.34		.50	
Girls	49	.35		.21		.20		.22		.15		-.07		.22		.29		.43		.42	
Boys and Girls	71	.32		.17		.29		.31		.19		.10		.32		.16		.41		.45	
Newton and Worcester																					
Boys	62	.59	.08	.42	.10	.51	.09	.50	.09	.45	.10	.33	.11	.41	.10	.39	.10	.49	.09	.60	.08
Girls	84	.30	.10	.34	.10	.23	.10	.16	.11	.15	.11	.18	.10	.22	.10	.20	.10	.33	.10	.33	.10
Boys and Girls	146	.42	.07	.23	.07	.36	.07	.29	.07	.29	.07	.24	.07	.29	.07	.25	.07	.38	.07	.45	.07

Table 25. Correlation Coefficients between Differential Aptitude Test Raw Scores and Teachers' Final Marks in English (Reduced to Standard Scores), and between Otis DM IQ's and Teachers' Final Marks in English (Reduced to Standard Scores), for All Groups Used in This Study

Group	N	Otis DM IQ		Verb. Reas.		Num. Abil.		Abst. Reas.		Space Rel.		Mech. Reas.		Cler. S & A		LU-I Spell.		LU-II Sent.		Total DAT	
		r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr	r	σr
Newton																					
Boys	40	.45		.54		.73		.49		.54		.05		.39		.60		.60		.67	
Girls	35	.47		.65		.39		.38		.22		.37		.39		.46		.48		.57	
Boys and Girls	75	.52		.59		.55		.44		.42		.01		.48		.53		.59		.62	
Worcester																					
Boys	22	.39		.12		.48		.18		.51		.14		.53		.22		.33		.46	
Girls	49	.42		.06		.15		.01		-.11		-.31		.07		.16		.36		.15	
Boys and Girls	71	.54		.09		.29		.08		.13		-.08		.32		.01		.38		.28	
Newton and Worcester																					
Boys	62	.75	.06	.48	.10	.68	.07	.45	.10	.56	.09	.22	.12	.57	.09	.36	.11	.56	.09	.66	.08
Girls	84	.50	.08	.45	.09	.34	.10	.25	.10	.13	.11	.14	.11	.23	.10	.29	.10	.47	.08	.41	.09
Boys and Girls	146	.55	.06	.27	.07	.45	.07	.28	,07	.29	.07	.04	.08	.41	.07	.30	.07	.51	.06	.48	.06

Test raw scores than between predictors and teachers' marks in general science. Correlations between predictors and teachers' marks in English are considerably higher than between predictors and teachers' marks in general science. The Otis DM IQ would seem to provide excellent prediction for the boys, as it correlates .75 with teachers' marks in English.

Summary of Results

In general, it may be said that the group used in this study was quite similar to the norm group used in the standardization of the Differential Aptitude Tests. Further, the Differential Aptitude Tests are fairly good predictors of success in the Read General Science Test, but are poor predictors for teachers' marks in general science and English. Finally, the Otis DM IQ is as good or a better predictor of success in most cases than any of the Differential Aptitude Tests.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Summary of Study and Implications of Results

Restatement of problem.-- The aim of this study was to obtain further validation data on the Differential Aptitude Tests in relation to success or failure in English and general science as offered in the ninth grade of the Worcester and Newton junior high schools.

Summary of findings.-- The study was run on 62 boys and 84 girls of the Bigelow Junior High School, Newton, Massachusetts, and several Worcester, Massachusetts, junior high schools. Differential Aptitude Test results were correlated with Read General Scoence Test raw scores and teachers' marks (reduced to Standard Scores) in English and general science. For the sake of comparison, Otis DM IQ's were also correlated with the Read Test and with teachers' marks. The results may be summarized as follows:

1. The study group compared quite favorably with the group used in determining the norms for the Differential Aptitude Tests. With the exception of Space Relations and Mechanical Reasoning, the mean of the boys in the study group was slightly above that of the norm group in each case. The study group, how-

ever, tended to be somewhat more homogeneous than the norm group. With the exception of Space Relations, the mean of the girls in the study group was slightly below that of the norm group in each case, although the study group tended to be somewhat more heterogeneous than the norm group.

2. Some of the intercorrelation coefficients of the Differential Aptitude Tests covered a wide range for the various groups. This may be accounted for by the low N's for some groups. Generally speaking, the coefficients for both sexes tended to be fairly close to those found for the norm groups.
3. The correlation coefficients between the Differential Aptitude Test raw scores and the Read General Science Test raw scores show a slight increase when combined. R's of .61 - .67 for boys and .62 - .71 for girls are obtainable by use of the correct pairs of Differential Aptitude Tests.
4. Correlation coefficients between Differential Aptitude Test raw scores and teachers' final marks in general science show that prediction is much poorer than for results on the Read Test. Considerably higher correlations are found with English grades.
5. Correlation coefficients between Otis DM IQ's and teachers' final marks in English and general science

and Read General Science Test raw scores are as high or higher than between Differential Aptitude Test raw scores and the above criteria.

Implications.-- The general conclusion seems to be that the differential Aptitude Tests, whether used singly or in combination, are no better predictors of success in ninth grade English or general science than is a general intelligence test of the Otis type. It is also apparent that little improvement in prediction may be had by selecting multiple predictors from among the Differential Aptitude Tests.

1. It was assured that teachers marked uniformly and accurately, which undoubtedly was not the case. This was partially compensated for by reducing teachers' marks to standard scores.
2. It was assured that the quantity and quality of teaching were the same for each group.
3. There are numerous other limitations: variability of motivation, varying absences on the part of different students, varying socio-economic factors, and varying extra-curricular activities, to mention a few.

Suggestions for Further Study

There are several possibilities for further study, not only problems stemming from this study, but also further

work on the Differential Aptitude Tests themselves.

1. Further validation of the Differential Aptitude Tests should be made in many more subjects and with larger N's.
2. The possibility of redesigning the Differential Aptitude Tests deserves consideration. No consistently high prognostic values have been obtained with the test in its present form.
3. Using the test results of this study, a follow-up study might be made of the students who continued with physics or chemistry courses. Extensions of Differential Aptitude Test data might be more fruitful than short-range data.

Table I. Central Tendency and Variability of Ninth Grade Students of the Bigelow Junior High School, Newton, Massachusetts

Tests		Boys			Girls			Boys & Girls		
		Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
Otis IQ...		106.73	12.12	33	109.78	10.47	27	107.75	11.97	60
DAT Raw Scores	Verb. Reas....	20.62	8.88	40	21.86	9.65	35	21.20	8.95	75
	Num. Abil....	19.50	9.88	40	19.71	8.14	35	19.60	9.10	75
	Abst. Reas....	27.38	9.18	40	28.86	9.72	35	28.05	9.50	75
	Space Rel.....	39.25	23.31	40	40.57	24.02	35	39.54	23.60	75
	Mech. Reas....	31.00	13.36	40	22.14	9.22	35	26.87	11.15	75
	Cler-ical....	49.12	12.14	40	56.86	8.98	35	52.73	11.40	75
	LU-I Spell...	34.50	25.76	40	46.00	22.95	35	40.36	24.75	75
	LU-II Sent....	26.62	13.39	40	36.57	15.74	35	31.27	15.60	75
Total DAT Score.....		247.50	88.26	40	271.07	85.32	35	258.50	88.40	75
Read Raw Score.....		43.75	10.58	40	37.86	11.05	35	41.00	11.20	75
English Mark.....		50.00	10.00	40	50.00	10.00	35	50.00	10.00	75
Science Mark.....		50.00	10.00	40	50.00	10.00	35	50.00	10.00	75

APPENDIX

Table II. Central Tendency and Variability of Ninth Grade Students of Several Worcester, Massachusetts, Junior High Schools

Tests		Boys			Girls			Boys & Girls		
		Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
Otis IQ...		107.05	8.41	22	100.58	10.34	49	102.60	10.67	71
DAT Raw Scores	Verb. Reas....	18.59	6.98	22	14.14	5.89	49	15.52	6.58	71
	Num. Abil....	13.36	7.10	22	12.71	6.06	49	12.92	6.41	71
	Abst. Reas....	22.00	10.44	22	18.84	11.19	49	19.82	11.06	71
	Space Rel.....	32.22	21.07	22	28.02	17.02	49	29.32	18.56	71
	Mech. Reas....	31.32	9.20	22	18.94	8.91	49	22.29	12.22	71
	Cler-ical....	42.91	11.24	22	49.45	13.02	49	47.42	12.86	71
	LU-I Spell...	34.50	16.70	22	30.78	19.31	49	32.50	18.92	71
	LU-II Sent....	24.27	8.49	22	23.64	13.12	49	23.83	11.88	71
Total DAT Score.....		218.68	53.14	22	195.40	55.76	49	202.62	56.02	71
Read Raw Score.....		40.50	9.74	22	33.82	8.23	49	35.88	9.26	71
English Mark.....		50.00	10.00	22	50.00	10.00	49	50.00	10.00	71
Science Mark.....		50.00	10.00	22	50.00	10.00	49	50.00	10.00	71

Table III. Central Tendency and Variability of Ninth Grade Students of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Tests		Boys			Girls			Boys & Girls		
		Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
Otis IQ...		106.51	11.31	55	103.84	11.29	76	104.96	11.38	131
DAT Raw Scores	Verb. Reas....	19.90	8.30	62	17.24	8.20	84	18.92	7.87	146
	Num. Abil....	17.32	9.46	62	15.63	7.80	84	16.35	8.59	146
	Abst. Reas....	25.47	9.98	62	23.01	11.70	84	24.06	11.07	146
	Space Rel.....	36.76	23.16	62	33.13	21.31	84	34.67	22.22	146
	Mech. Reas....	30.71	12.08	62	19.68	9.68	84	24.36	12.07	146
	Cler- ical....	46.92	12.20	62	52.54	12.08	84	50.15	12.44	146
	LU-I Spell...	34.50	22.96	62	37.60	22.38	84	36.59	22.44	146
	LU-II Sent....	25.79	11.94	62	29.02	15.70	84	27.65	14.26	146
Total DAT Score.....		237.60	78.82	62	226.92	79.06	84	231.32	79.14	146
Read Raw Score.....		42.60	10.40	62	35.50	10.45	84	38.52	10.61	146
English Mark.....		50.00	10.00	62	50.00	10.00	84	50.00	10.00	146
Science Mark.....		50.00	10.00	62	50.00	10.00	84	50.00	10.00	146

Table IV. Correlation Coefficients for 40 Ninth Grade Boys of the Bigelow Junior High School, Newton, Massachusetts

Test / Test	Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ*	.63	.45	.53	.87	.47	.41	.47	.38	.17	.72	.46	.36
DAT Raw Scores: Verb. Reas.	.59	.54	.58	.70	.51	.56	.22	.32	.62	.60	.57	
DAT Raw Scores: Num. Abil.	.63	.73	.48	.74	.56	.64	.46	.17	.53	.71		
DAT Raw Scores: Abst. Reas.	.54	.49	.76	.88	.76	.60	.73	.46	.62			
DAT Raw Scores: Space Rel.	.60	.54	.48	.82	.46	.62	.30	.45				
DAT Raw Scores: Mech. Reas.	.28	.05	.33	.54	.22	.08	-.06					
DAT Raw Scores: Clerical	.46	.39	.28	.50	.11	.19						
DAT: LU-I Spell.	.53	.60	.54	.78	.63							
DAT: LU-II Sent.	.58	.60	.62	.73								
Total DAT Score	.69	.67	.71									
Read Raw Score	.66	.56										
English Mark	.78											

*N = 33

Table V. Correlation Coefficients for 35 Ninth Grade Girls of the Bigelow Junior High School, Newton, Massachusetts

Test \ Test		Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell.	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ*		.58	.47	.38	.92	.59	.49	.48	.61	.60	.68	.38	.70
DAT Raw Scores	Verb. Reas.	.79	.65	.74	.76	.78	.68	.56	.46	.56	.53	.50	
	Num. Abil.	.64	.39	.53	.70	.60	.46	.30	.31	.44	.47		
	Abst. Reas.	.41	.38	.37	.67	.53	.32	.29	.37	.55			
	Space Rel.	.33	.22	.52	.78	.34	.32	.35	.60				
	Mech. Reas.	.34	.37	.46	.66	.58	.34	.18					
	Clerical	.54	.39	.32	.62	.48	.43						
	LU-I Spell.	.49	.46	.46	.74	.65							
	LU-II Sent.	.55	.48	.60	.88								
Total DAT Score		.67	.57	.70									
Read Raw Score		.62	.46										
English Mark		.68											

*N = 27

Table VI. Correlation Coefficients for 75 Ninth Grade Boys and Girls of the Bigelow Junior High School, Newton, Massachusetts

Test / Test		Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell.	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ*		.64	.52	.45	.95	.57	.49	.61	.50	.52	.92	.48	.55
DAT Raw Scores	Verb. Reas.	.68	.59	.64	.81	.65	.59	.37	.37	.64	.59	.69	
	Num. Abil.	.62	.55	.48	.71	.52	.52	.38	.22	.52	.60		
	Abst. Reas.	.46	.44	.53	.78	.62	.43	.55	.40	.63			
	Space Rel.	.51	.42	.59	.83	.54	.45	.37	.54				
	Mech. Reas.	.36	-.01	.52	.46	.22	.03	-.12					
	Clerical	.35	.48	.18	.55	.34	.25						
	LU-I Spell.	.42	.53	.35	.70	.62							
	LU-II Sent.	.46	.59	.46	.78								
Total DAT Score		.64	.62	.63									
Read Raw Score		.65	.35										
English Mark		.61											

*N = 60

Table VII. Correlation Coefficients for 22 Ninth Grade Boys of several Worcester, Massachusetts, Junior High Schools

Test / Test		Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT RAW SCORES: LU-II Sent.	LU-I Spell.	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ		.34	.39	.72	.47	.64	-.42	.13	.21	.58	.17	.10	-.02
DAT Raw Scores	Verb. Reas.	.13	.12	.48	.34	.59	-.03	-.16	.47	.10	.13	.08	
	Num. Abil.	.42	.48	.03	.43	.14	.06	.20	.24	.31	.22		
	Abst. Reas.	.47	.18	.27	.49	.40	-.13	.22	.39	.28			
	Space Rel.	.24	.51	.29	.63	.41	-.23	.31	.21				
	Mech. Reas.	.18	.14	.29	.45	.57	-.34	.18					
	Clerical	.50	.53	-.01	.28	.18	-.20						
	LU-I Spell.	.09	-.22	.18	.08	.16							
	LU-II Sent.	.34	.33	.58	.71								
Total DAT Score		.50	.46	.47									
Read Raw Score		.26	.06										
English Mark		.54											

Table VIII. Correlation Coefficients for 49 Ninth Grade Girls of Several Worcester, Massachusetts, Junior High Schools

Test / Test	Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ*	.35	.42	.73	.55	.77	.63	-.39	.35	.21	-.06	.18	.29
DAT Raw Scores: Verb. Reas.	.21	.06	.50	.49	.53	.33	-.16	.14	.23	.25	.12	
DAT Raw Scores: Num. Abil.	.20	.15	.21	.34	.37	.17	-.16	-.06	.20	.18		
DAT Raw Scores: Abst. Reas.	.22	.01	.38	.63	.39	.17	.01	.34	.49			
DAT Raw Scores: Space Rel.	.15	-.11	.30	.55	.19	-.08	.21	.16				
DAT Raw Scores: Mech. Reas.	-.07	-.31	-.07	.25	.06	.03	-.31					
DAT Raw Scores: Clerical	.22	.07	-.32	.13	-.01	-.16						
DAT Raw Scores: LU-I Spell.	.29	.16	.53	.58	.60							
DAT Raw Scores: LU-II Sent.	.43	.36	.68	.75								
Total DAT Score	.42	.15	.56									
Read Raw Score	.32	.27										
English Mark	.55											

Table IX. Correlation Coefficients for 71 Ninth Grade Boys and Girls of Several Worcester, Massachusetts, Junior High Schools

Test / Test		Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ		.32	.54	.71	.52	.69	.10	-.30	.39	.32	.03	.13	.25
DAT Raw Scores	Verb. Reas.	.17	.09	.55	.46	.51	.16	-.22	.42	.20	.24	.11	
	Num. Abil.	.29	.29	.14	.36	.30	.07	-.06	.13	.24	.18		
	Abst. Reas.	.31	.08	.37	.60	.39	.05	.03	.40	.42			
	Space Rela.	.19	.13	.31	.58	.24	-.16	.15	.24				
	Mech. Reas.	.10	-.08	.35	.46	.22	.01	-.08					
	Clerical	.32	.32	-.28	.12	.03	-.35						
	LU-I Spell.	.16	.01	.30	.34	.44							
	LU-II Sent.	.41	.38	.61	.72								
Total DAT Score		.45	.28	.55									
Read Raw Score		.29	.21										
English Mark		.61											

Table X. Correlation Coefficients for 62 Ninth Grade Boys of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Test / Test	Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ*	.59	.75	.62	.74	.54	.28	.44	.31	.48	.72	.40	.33
DAT Raw Scores: Verb. Reas.	.42	.48	.56	.62	.53	.42	.14	.43	.47	.46	.46	
DAT Raw Scores: Num. Abil.	.51	.68	.38	.69	.47	.50	.43	.23	.48	.58		
DAT Raw Scores: Abst. Reas.	.50	.45	.59	.74	.63	.36	.57	.50	.50			
DAT Raw Scores: Space Rel.	.45	.56	.43	.75	.44	.40	.32	.07				
DAT Raw Scores: Mech. Reas.	.33	.22	.50	.50	.36	.04	.13					
DAT Raw Scores: Clerical	.41	.57	.21	.45	.14	.04						
DAT Raw Scores: LU-I Spell.	.39	.36	.44	.64	.54							
DAT Raw Scores: LU-II Sent.	.49	.56	.60	.72								
Total DAT Score	.60	.66	.64									
Read Raw Score	.50	.49										
English Mark	.75											

*N = 55

Table XI. Correlation Coefficients for 84 Ninth Grade Girls of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Test \ Test		Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ*		.30	.50	.83	.77	.75	.42	-.23	.35	.45	.31	.38	.57
DAT Raw Scores	Verb. Reas.	.34	.45	.69	.87	.78	.59	.33	.50	.56	.54	.68	
	Num. Abil.	.23	.34	.40	.64	.58	.37	.15	.29	.43	.43		
	Abst. Reas.	.16	.25	.38	.70	.54	.30	.20	.48	.57			
	Space Rel.	.15	.13	.44	.73	.49	.20	.29	.49				
	Mech. Reas.	.18	.14	.40	.63	.44	.27	.19					
	Clerical	.22	.23	.01	.41	.27	.06						
	LU-I Spell.	.20	.29	.41	.64	.63							
	LU-II Sent.	.33	.47	.61	.85								
Total DAT Score		.33	.41	.60									
Read Raw Score		.34	.35										
English Mark		.48											

*N = 76

Table XII. Correlation Coefficients for 146 Ninth Grade Boys and Girls of the Bigelow Junior High School, Newton, Massachusetts, and Several Worcester, Massachusetts, Junior High Schools

Test / Test		Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT LU-II Sent.	DAT LU-II Spell.	RAW Clerical	RAW Mech. Reas.	SCORES Space Rel.	SCORES Abst. Reas.	SCORES Num. Abil.	SCORES Verb. Reas.
Otis IQ*		.42	.55	.72	.79	.66	.34	.14	.38	.47	.47	.39	.47
Scores	Verb. Reas.	.23	.27	.45	.77	.56	.39	-.06	.37	.44	.39	.48	
Scores	Num. Abil.	.36	.45	.41	.66	.51	.47	.25	.28	.45	.49		
Scores	Abst. Reas.	.29	.28	.48	.60	.55	.28	.31	.48	.54			
Scores	Space Rel.	.29	.29	.45	.74	.45	.26	.28	.44				
Raw	Mech. Reas.	.24	.04	.54	.54	.30	.08	.04					
Raw	Clerical	.29	.41	.02	.41	.24	.01						
DAT	LU-I Spell.	.25	.30	.35	.60	.58							
DAT	LU-II Sent.	.38	.51	.55	.79								
Total DAT Score		.45	.48	.63									
Read Raw Score		.41	.30										
English Mark		.57											

*N = 131

Table XIII. Median Correlation Coefficients for Total Nine Sub-groups of the Study

Test / Test		Science Mark	English Mark	Read Raw Score	Total DAT Score	DAT Raw Scores: LU-II Sent.	LU-I Spell.	Clerical	Mech. Reas.	Space Rel.	Abst. Reas.	Num. Abil.	Verb. Reas.
Otis IQ		.42	.50	.71	.77	.64	.42	.39	.38	.47	.42	.38	.36
DAT Raw Scores	Verb. Reas.	.34	.45	.56	.70	.56	.42	.14	.42	.47	.46	.48	
	Num. Abil.	.42	.48	.40	.66	.51	.46	.25	.23	.45	.47		
	Abst. Reas.	.41	.28	.38	.67	.54	.30	.22	.40	.54			
	Space Rel.	.29	.29	.44	.74	.44	.26	.30	.44				
	Mech. Reas.	.24	.05	.40	.50	.30	.04	.12					
	Clerical	.35	.39	.02	.41	.18	.04						
	LU-I Spell.	.29	.30	.41	.64	.60							
	LU-II Sent.	.43	.48	.61	.75								
Total DAT Score		.50	.48	.63									
Read Raw Score		.41	.35										
English Mark		.61											

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Bear, R.M., "The Predictive Value of the Iowa Physics Aptitude Test", Journal of Applied Psychology, Vol. II, No. 10, October, 1927, pp. 381 - 384.

2. Bennett, G.K., H.G. Seashore, and A.G. Wesman, Differential Aptitude Tests, The Psychological Corporation, New York, 1947.

3. Benton, A.L., and J.D. Perry, "A Study of the Predictive Value of the Stanford Scientific Aptitude Test (Zyve)", Journal of Psychology, Vol. 10, 1940, pp. 309 - 312.

4. Bingham, W.V., Aptitudes and Aptitude Testing, Harper & Brothers, New York, 1937.

5. Bittner, R.H., and G.E. Wilder, "Expectancy Tables: A Method of Interpreting Correlation Coefficients", Journal of Experimental Education, March, 1946, pp. 245 - 252.

6. Bridges, C.F., "Some Basic Considerations in Determining the Significance of Achievement Test Results", New Jersey Elementary Principals' Bulletin, Vol. 21, No. 6, December, 1948.

7. Burnham, P.S., and A.B. Crawford, Forecasting College Achievement - Part I, Yale University Press, New Haven, Connecticut, 1946.

8. Challman, R.C., A.I. Gates, A.T. Jersild, and T.R. McConnell, Educational Psychology, MacMillan Company, New York, 1950.

9. Chittendon, E.W., C.J. Lapp, and D.B. Stuit, Physical Science Aptitude Examination in Special Education, State University of Iowa, Iowa City, Iowa, 1943.

10. Cottle, W.G., "The Differential Aptitude Tests: Some Comments", Occupations, Vol. 26, No. 6, March, 1948, pp. 344-- 345.

11. Crawford, A.B., "Stanford Scientific Aptitude Test (Zyve)", Mental Measurements Yearbook, The Mental Measurements Yearbook, Highland Park, New Jersey, 1940, pp. 453 - 454.

12. Crawford, A.B., "What About All These Tests?", Occupations, Vol. 12, No. 8, April, 1934, pp. 13 - 18.

13. Doppelt, J.E., and H.G. Seashore, "How Effective Are Your Tests?", Test Service Bulletin, No. 37, June, 1949, The Psychological Corporation, New York.

14. Dubins, M.I., A Review of the Research in Testing in Secondary Chemistry and Physics from 1938 - 1948 Including 4 Reviews of Standardized Tests, Unpublished Ed. M. Thesis, Boston University School of Education, 1949.

15. Durost, W.N., The Measurement Notebook, Boston University, Boston, Massachusetts, 1950.

16. Durost, W.N., and H.M. Walker, Durost - Walker Correlation Chart, World Book Company, New York, 1937.

17. Fromm, E., Man for Himself, Rinehart & Company, Incorporated, New York, 1941.

18. Garrett, H.E., "A Developmental Theory of Intelligence", American Psychologist, Vol. 1, No. 9, September, 1946.

19. Good, C.V., Dictionary of Education, McGraw - Hill Book Company, New York, 1945, p. 28.

20. Hildreth, G.H., A Bibliography of Mental Tests and Rating Scales, The Psychological Corporation, New York, 1939.

21. Hull, C.L., "Variability in Amount of Different Traits Possessed by the Individual", Journal of Educational Psychology, Vol. 18, No. 2, February, 1927, pp. 97 - 106.

22. Hull, C.L., and C.E. Limp, "The Differentiation of the Aptitudes by Means of Test Batteries", Journal of Educational Psychology, Vol. 16, No. 2, February, 1925, pp. 73 - 88.

23. Kilzer, L.R., and T.J. Kirby, Kilzer - Kirby Inventory Test for Mathematics Needed in High School Physics, Public School Publishing Company, Bloomington, Illinois, 1929.

24. Kvaraceus, W.C., Methods in Educational Research, Boston University, Boston, Massachusetts, 1949.

25. Lapp, C.J., and D.B. Stuit, "Some Factors in Physics Achievement at the College Level", Journal of Experimental Education, Vol. 9, No. 3, March, 1941, pp. 251 - 253.

26. Lindquist, E.F., A First Course in Statistics, Houghton Mifflin Company, Cambridge, Massachusetts, 1942.

27. Lindquist, E.F., Educational Measurement, American Council on Education, Washington, D.C., 1951.

28. Monroe, W.S., Encyclopedia of Educational Research, MacMillan Company, New York, 1941.

29. Mursell, J.L., Psychological Testing, Longmans, Green & Company, New York, 1950.

30. Peatman, J.G., Descriptive and Sampling Statistics, Harper & Brothers, New York, 1947.

31. Ricks, J.H., Jr., and H.G. Seashore, "Norms Must Be Relevant", Test Service Bulletin, No. 39, May, 1950, The Psychological Corporation, New York.

32. Ruch, F.L., Psychology and Life, Scott, Foresman & Company, New York, 1941.

33. Segel, D., "The Validity of a Multiple Aptitude Test at the Secondary School Level", Educational and Psychological Measurement, Vol. 7, No. 4, Winter, 1947, pp. 675 - 705.

34. Super, D.E., "The Place of Aptitude Testing in the Public Schools", Educational and Psychological Measurement, Vol. 2, No. 3, Summer, 1942, pp. 267 - 277.

35. Thorndike, E.L., The Psychology of Learning, Teachers' College, Columbia University, New York, 1913.

36. Traxler, A.E., "Evaluation of Aptitude and Achievement in a Guidance Program", Educational and Psychological Measurement, Vol. 6, No. 1, Spring, 1946, pp. 3 - 16.

37. Wesman, A.G., "Expectancy Tables - A Way of Interpreting Test Validity", Test Service Bulletin, No. 38, December, 1949, The Psychological Corporation, New York.

38. Wesman, A.G., "The Three-Legged Coefficient", Test Service Bulletin, No. 40, December, 1950, The Psychological Corporation, New York.

39. Wesman, A.G., "What is an Aptitude?", Test Service Bulletin, No. 36, August, 1948, The Psychological Corporation, New York.

40. Yagjian, C., A Validation Study of the Differential Aptitude Tests in Relation to First Year High School Vocational Bookkeeping in Two Communities, Unpublished Ed. M. Thesis, Boston University School of Education, Boston, Massachusetts, 1950.

41. Zyve, D.L., "A Test of Scientific Aptitudes", Journal of Educational Psychology, Vol. 18, No. 8, August, 1927, pp. 525 - 546.

42. Zyve, D.L., Stanford Scientific Aptitude Test, Stanford University, Palo Alto, California, 1927.