Educational guidance, an experimental study in the analysis and prediction of ability of high school pupils, by Truman Lee Kelley.

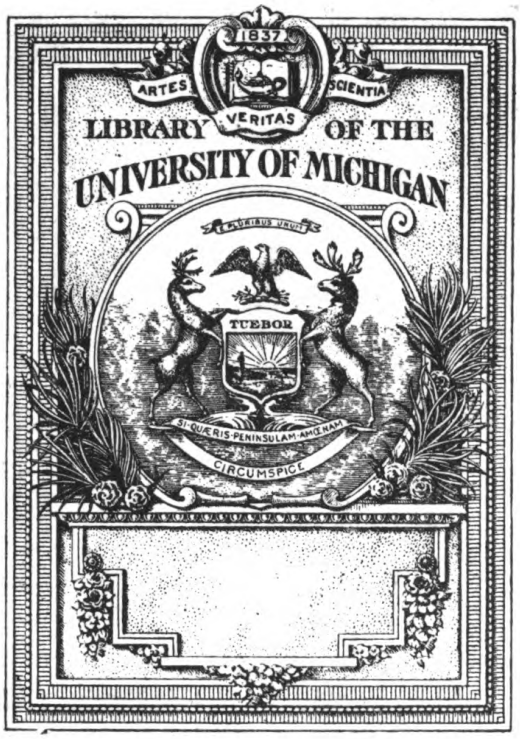
Kelley, Truman Lee, 1884-

New York, Teachers College, Columbia University, 1914.

Columbia University

Contributions to Education

Teachers College Series



EDUCATIONAL GUIDANCE

AN EXPERIMENTAL STUDY IN THE ANALYSIS AND PREDICTION OF ABILITY OF HIGH SCHOOL PUPILS

BY

TRUMAN LEE KELLEY, PH.D.

CONSULTING PYSCHOLOGIST, CULVER MILITARY ACADEMY

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
CONTRIBUTIONS TO EDUCATION, NO. 71

PUBLISHED BY

Teachers College, Columbia University

NEW YORK CITY

1914

PREFACE

The task of giving tests, establishing averages, and calculating relations, which shall serve as a basis for prognosis of mental ability, is, in every sense, a social undertaking, and it is only because of the kindly coöperation of the principals, teachers, and pupils of the two schools studied that it has been possible to secure the data that supply the material for this investigation. The evaluation of the data has equally been a social task and I am particularly indebted to Professors E. L. Thorndike, S. S. Colvin, and H. A. Ruger for assistance in grading the preferences of pupils in the interest test as to vocations, sports, and entertainments, and to Mrs. Grace Osgood and Miss Grace Kelley for the unique assistance which, as librarians, they were able to render in grading magazines and books.

To the ever-ready, stimulating criticisms of Professor Thorndike, I am peculiarly indebted, for it is due to his encouragement that the investigation covers the three fields of mathematics, English, and history instead of one only, and that the number of relations determined is as extensive as it is. The field covered gives the work whatever of value it has, but the accomplishment of it and its appearance in print at this time has been possible only because of the devoted and untiring assistance, in grading, calculating coefficients of correlation, and deriving regression equations, rendered by my wife.

September, 1914. T. L. KELLEY.

CONTENTS

EDUCATIONAL GUIDANCE

SECTION 1.—STATEMENT OF THE PROBLEM

Humanity's unvoiced plea for guidance is the foundation of all professions. The doctor, the lawyer, the minister find that belief and obedience are more often the result of need than of understanding and conviction. The modern idea of education is crystallizing into an effort to guide rather than to instruct—to answer to a need rather than to cater to a curriculum. The growing recognition of the need for vocational and educational guidance is resulting in the establishment of bureaus endeavoring to give the former, and in the training of psychologists to solve the problems of the latter.

The movement for vocational guidance is in its infancy, but it only depends upon improved methods and more extended research to give it a place with the older professions. Vocational guidance has sprung up out of two needs—the need of the employer for efficient clerks, mechanics, and laborers and, still more important, the need of the individual to utilize his talents to the best advantage in order to cope with present-day industrial conditions.

This latter demand is most pressing at the time that the individual is about to leave school, and it is at this point that the major efforts of vocational guidance bureaus have been expended; but even a hasty consideration will show that the guidance exercised is tardy. It should have been present when the school training of the individual became different from that of other individuals—when he began to specialize and train himself for his life work. It may be stated with assurance that in all cases this specialization should be well under way before the completion of the formal education of the pupil.

These remarks suffice to make apparent the need for such educational guidance in the high school and college, as shall precede and serve as a basis for the later vocational guidance.

The general method to determine the accuracy of guidance is the same, whether the guidance be educational or vocational, and it is one of the chief aims of this study to determine accurately the reliability of the estimation of academic capacity. The data necessary for accomplishing this are at hand, for high school records of academic accomplishment are universally kept. Comparable vocational records are generally not available; but for the determination of the reliability of an estimate of vocational fitness they are essential, and whenever available the method here used is applicable.

The two chief factors entering into the problem of efficient guidance are, first, a correct understanding of the demands of prospective tasks and, second, an accurate valuation of the ability of the person in question to meet these demands. These two main elements of the problem may be stated as requiring an analysis of the individual to determine his characteristics, and an analysis of the needs of the situation to see to what extent the individual meets these needs. This is a general statement of the problem applicable to all kinds of guidance. The problem here undertaken is termed one in educational guidance, since the data concern high school pupils and high school subjects; but the method, which is that of calculating the correlation between the estimate of a person's fitness for a task and his later performance in it, is of general validity and importance and will inevitably be used extensively in vocational guidance.

As success usually depends upon several factors, partial correlation and the regression equation method are essential in the evaluation of the data. This method will be explained more fully later. The writer is not aware that it has been used before in a guidance problem, but its peculiar adaptability to a problem of this nature insures its extended use in the future.

More specifically, the endeavor of this study is to predict with a known, and as high as possible, degree of accuracy the capacity of the pupil to carry a prospective high school course. In doing this, an analysis of the factors which make for success in the course is obtained. The essential objects of the study are thus (1) a measurement of the characteristics of the pupil, together with the determination of the extent to which these characteristics correlate with scholastic ability along certain lines, and (2) an analysis of the demands of certain high school courses.

To illustrate the intimacy of these two problems it may be pointed out that if all the essentials of fitness needed to fulfill a certain task were known, and if the abilities of the person under consideration were completely known, then prediction and performance would agree perfectly; and to the extent that this condition is approximated, the correlation between prediction and performance is increased.

SECTION 2.—METHOD AND SPECIFIC OBJECT

When selective classification of a prospective high school pupil is attempted, the usual question asked is, what is his general mental ability, and he is classified according to the answer to that question. The present study attempts to answer that question by considerations based upon one of three sources of data: (1) the pupil's grammar school record, (2) estimates of previous teachers of the pupil, and (3) grades obtained in special tests given the pupil at the very beginning of the school year. Beyond this, it is imperative, in rendering the most valid decision as to the pupil's capacities, that account be taken of his specific interests and peculiar genius. An excellent student of mathematics may be a very poor English scholar, and though this situation is not true in the majority of cases, yet the number of cases in which it is true is sufficiently great that very material injustice will be worked if it is not taken into consideration.

The further aim of this study is, therefore, to determine, before courses in the high school are taken, what the probable ability of the pupil in question will be in them. Instead of attempting to cover the field of high school work exhaustively, three subjects —mathematics, English and history—have been selected for study. The general method of procedure with all three subjects and all three sources of data is to separate the data into elements that are, as far as possible, independent of each other, e.g., the teachers' estimates of the pupil are four in number, (1) intellectual ability, (2) conscientiousness, (3) emotional interest in his work, and (4) oral expression. All of these factors are important for scholastic work and it would be desirable if they were totally uncorrelated with each other. The first and fourth and the second and third are rather closely related with each other, but even so there is sufficient independence between the four to make their combined significance as indicators of scholastic success considerably greater than that of a single estimate, such as that of intellectual ability.

If the grades received, or marks given, in the original data are represented by X_1, X_2, X_3, X_4, and if the grades received in the high school mathematics, English and history courses six months

or a year after the original data are obtainable are represented by X_M, X_E, X_H, then the problem is to establish the correlation between X_M and the combined measures based upon X_1, X_2, X_3, X_4, and similarly with X_E and X_H. Expressed as an equation it is $X_M = c_0 + c_1X_1 + c_2X_2 + c_3X_3 + c_4X_4$. This is equivalent to saying that a certain constant times the grade received in the first trait (or test), plus a second constant times the grade received in the second trait (or test), plus, etc., gives the probable grade in the course about to be taken. The statistical problem involved is the determination of the constants c_0, c_1, c_2, c_3, c_4, so that the X_M values obtained differ on the whole, and when every individual is taken into account, from the actual X_M values by the smallest amount possible.[1]

The equation which fulfills this condition is called a regression equation, and the constants c_1, c_2, c_3, c_4, are called regression coefficients. They are functions of the coefficients of correlation between the various X's and the standard deviations of the X's. The theoretical proof of the derivation of these constants may be found in Yule, "Introduction to the Theory of Statistics," and a considerable amount of the purely mathematical work involved in their calculation is given in the Appendix of the present work.[2] For an understanding of this investigation (except the Appendix) and the use of the method, it will suffice if the reader has well in mind the fact that the value X_M for each individual obtained by this equation is the most probable value which it is possible to obtain from the data X_1, X_2, X_3, X_4[3].

This regression equation is the means of prognosis, and to use it in the case of any individual it is only necessary to substitute the values X_1, X_2, X_3, X_4, for that individual, to obtain a value X_M.

In addition to knowing the value X_M, it is essential to know the probable error of it, or to know its standard deviation. This has been calculated in all cases, that the reliability of the prognosis may be known. This reliability depends upon two factors, the reliabilities of X_1, X_2, X_3, X_4, and the extent to which these

[1] Or, more accurately, that the calculated X_M's differ from the actual X_M's by such amounts that the sum of the squares of the differences is a minimum.

[2] The writer is about to publish tables which will greatly facilitate the calculation of regression equations.

[3] For the mathematician the words "in case the regression is rectilinear" may be added.

X's are correlated with X_M (this latter is in part dependent upon the former). The reliability of any given measure X_1 is given by the reliability coefficient,[1] which is simply the value of the coefficient of correlation between the given set of X_1's and a second set similarly derived. To obtain this measure it is necessary to have the X_1 grades assigned by at least two judges, which procedure has been followed throughout except where impossible because of the nature of the data, or where totally unnecessary because the grading was so completely defined that the judge had little or no option left to him in his grading. The formula giving this reliability coefficient of a grade, which is the average or sum of the gradings of n judges, is $\frac{nr}{1+(n-1)r}$ where r is the correlation between gradings of different judges. Most of the tests in this study have been graded by two judges, so that the formula becomes $\frac{2r}{1+r}$.

It is later explained at some length that the use of correlation coefficients, corrected for attenuation, is not permissible in this problem. The attempt here is to prophesy accomplishment by measuring an existing, not an imaginary, relationship, whereas, in the studies using methods for "correcting" raw coefficients of correlation, the attempt is to obtain a coefficient which is an estimate of an ideal relationship and which does not represent a correlation between existent data. This distinction should be clearly borne in mind and comparison should not be made with studies using coefficients corrected for attenuation.

In addition to being the means of prognosis, the regression equation serves one other important function: the regression coefficient c_1 gives the weight that must be attached to the measure X_1, independent of and free from any relation it may have with X_2, X_3, X_4. It therefore makes it possible to consider the importance of each of the factors X_1, X_2, X_3, X_4, independent of the others. Such an analysis is essential in arriving at the separate factors which go to make up efficiency in any given subject. This latter use of the coefficients of the regression equation will be more apparent when treating of teachers' estimates and the special tests, than in the following section covering the use of elementary school grades as indicators of high school ability.

[1] See Brown, Mental Measurement, pp. 101–102.

SECTION 3.—SCHOOL GRADES AND THEIR SIGNIFICANCE AS EVIDENCE OF HIGH SCHOOL EFFICIENCY

The data here treated consist of the scholastic records of 59 pupils whose grades were available from the fourth grade through the first year of the high school. These pupils had attended the same school without a break, except for minor illness in certain cases, during this period. The grades of the pupils in the following subjects were copied from the high school records: FA (first-year average grade), FM (first-year mathematics-algebra), FE (first-year English),[1] 7A (7th grade average grade), 7M (7th grade mathematics-arithmetic), 7E (7th grade English), 7H (7th grade history), 6A, 6M, 6E, 6H, 5A, 5M, 5E, 5H, 4A, 4M, 4E, 4H.[2]

The coefficients of reliability of these measures are not available, but they are probably not less than .80 for FA, 6A, 4A, and not less than .75 for 7A and 5A. It was first determined what connection there is between 7A, 6A, 5A, 4A, and FA. In order to determine this the correlation between each one of these grades and all the rest is necessary. These correlations are given in the following table:

	FA	7A	6A	5A
7A	.719			
6A	.728	.730		
5A	.531	.425	.541	
4A	.624	.551	.573	.576

There are several surprising items in this table. All of the correlations involving 6A and 4A are higher than would be expected from the balance of the data. It certainly would not be expected that 6th grade marks would correlate more highly with first year standing than 7th grade marks, nor that 4th grade marks would correlate more highly with first year and 7th grade marks than 5th grade marks. It is possible that the teachers of the 6th and 4th grades were more expert in estimating the ability

[1] See Appx., pp. 108–116.

[2] See Appx., p. 116.

of their pupils than were the teachers in the 7th and 5th grades. However this may be, to get the most out of these particular data in their bearing upon FA, the regression equation, based upon these coefficients of correlation and the various standard deviations, must be obtained. Calculation shows it to be as follows:[1] $C \cdot F_A = 1.67(7_A) + 1.3(6_A) + .4(5_A) + .7(4_A)$. (In which C is some constant.) Calculation shows the correlation between the FA's thus obtained, and the FA's actually obtained in the first year to be .789, with a probable error of .032. This correlation will be designated by the symbol $r_{F_A(7, 6, 5, 4_A)}$. This is a high correlation for data so far apart in time, and the division of pupils in the high school into sections according to ability, upon the basis of this prognosis, would be much more accurate than that which would be possible after observing the progress of the pupil in the high school for half a school year; for this correlation is undoubtedly higher than that between average half-year term grades. Especially would this be true if succeeding term grades were given by different instructors.

The argument that this correlation is not perfect and would work injustice in certain cases is utterly impotent if the alternative is the present very common system of mixing the good, the medium and the poor all together, thus actually doing injury to all. For any high school having more than a single section of each class, and where grammar grade records are available, the desirability of classification upon the basis here worked out will be apparent, whether considered from the standpoint of the nervous strain upon the teachers of a non-homogeneous class, from the standpoint of economical administration, or from the standpoint of the accomplishment of the pupil. In this connection it should be mentioned that the accuracy of a classification based upon the marks received in the 7th grade alone is not very materially less than that which is based upon the marks from the 4th to the 7th grades, and would be of very decided value in case more extended records are not available.

There is one drawback to the use of the above regression equation, viz: by its use that pupil, who is particularly capable in some one line, is not classified more highly in that line than he is in others. A more detailed estimate of ability is desirable and

[1] See Appx., p. 91.

can be obtained by calculating the regression equations to estimate ability in the various subjects of the first-year class, instead of one regression equation to estimate average high school ability.

The most probable grade in first-year mathematics (FM) would be determined from the grades received in the different elementary school subjects for the years for which the data are available, i.e., the most probable value of FM is equal to some combination of 7M (7th grade mathematics), 6M, 5M, 4M, and also 7E (7th grade English), 6E, 5E, 4E, and so forth for the balance of the elementary school curriculum. The grades in only three elementary school subjects, mathematics, English and history, were taken from the school records (it is the average of these three that give the grades 7A, 6A, 5A, 4A), since these subjects all run through the last four years of the elementary school, and since the means of the various grades for these subjects could be determined with considerable accuracy, probably very much greater accuracy than with such subjects as nature study, writing, music, etc. Furthermore, FM is undoubtedly more especially dependent upon the grades 7M, 6M, 5M, 4M, than upon grades in other subjects in the curriculum, and similarly with FE and 7E, etc. It may also be stated that for purposes of determining the difference of capacity of a pupil for mathematics and his capacity for English there is very little gained by involving a subject such as history in the calculation. For these reasons, the bearing of 7M, 6M, 5M, 4M only upon FM has been obtained, and in doing this it was assumed that the importance of the various years of the elementary school was the same as in the case of the average first year standing and the average standings of the elementary grades. The equation of relation (the term regression equation is reserved for equations satisfying entirely the conditions laid down for such equations) is therefore as follows:

$$C \cdot \text{FM} = 1.67(7\text{M}) + 1.3(6\text{M}) + .4(5\text{M}) + .7(4\text{M})$$

Similarly $$C \cdot \text{FE} = 1.67(7\text{E}) + 1.3(6\text{E}) + .4(5\text{E}) + .7(4\text{E})$$

The correlation between the FM's thus obtained and the actual FM's is .580 ($r_{\text{FM}\,(7,\,6,\,5,\,4\text{M})}$). For English $r_{\text{FE}\,(7,\,6,\,5,\,4\text{E})} = .710$. The greater correlation in the case of English than in the case of mathematics may be partly due to an intrinsic difference in the laws of development of an individual with reference to these two subjects, but it is, at least in part, due to the greater reliability

of the English elementary school marks, since these measures are an average of the grades given in two English courses, whereas the arithmetic grades are obtained from but a single course. It is evident that there is also a greater content difference in passing from arithmetic to algebra than in passing from 7th grade English to first-year English. From a statistical point of view it does not seem likely that the difference in reliability could entirely account for the difference in correlation, and the author will state that the mathematical probability of the difference being due to chance is small, though he cannot express this probability in exact numerical terms.

It has been stated that the value of these coefficients of correlation lies in their power to differentiate between the ability of the pupil in mathematics and in English. The extent to which they perform this task in differential diagnosis can be measured by comparing for each individual the difference between the estimated ability in mathematics and the estimated ability in English with the actual difference of ability as shown by the grades in the two subjects. If individual (1) is estimated to be .7 sigma (standard deviation) above the average in mathematics and .4 sigma above the average in English, and the actual grades which he received are .9 sigma above the average in mathematics and .6 sigma above the average in English, then the estimated difference between the abilities in the two subjects is equal to the actual difference.

The extent to which differences in estimation correspond to differences in first-year grades is given by the coefficient of correlation between these two differences, $r_{(\mathrm{FM-E})\ (7,\ 6,\ 5,\ 4,\ \mathrm{M-E})}$. It is evident that if this correlation equals zero, then English grades in the elementary school are as good a basis for estimation of mathematics grades in the first-year class of the high school as are mathematics grades in the elementary school—in other words, intelligence is general, and may be directed by the individual with equal result in any direction. On the other hand, if the correlation is perfect,[1] then mental capacity is specific and specialized to exactly the same extent and in the same manner

[1] For this theoretical consideration, not in the nature of a prognosis, a coefficient of correlation corrected for attenuation might be desired, but the data for its calculation are not available, nor is it likely that the assumptions underlying its derivation (lack of correlation of errors, etc.) would be sound. Such correction, if utilized, would increase the correlation found.

in the high school and in the elementary school when dealing with the same subjects. Calculation shows that $r_{(F_{M-E})\ (7,\,6,\,5,\,4,\,M-E)} = .515$.

The net conclusion which may be drawn from these four coefficients of correlation is, that it is possible to estimate a person's general ability in the first year class from the marks he has received in the last four years of the elementary school with an accuracy represented by a coefficient of correlation of .789; and that individual idiosyncracies may be estimated, in the case of mathematics and English, with an accuracy represented by a coefficient of correlation of .515.

The method of doing this is the simple one of substitution in a regression equation. The regression equation given above proved the best for the school from which the data are obtained, but it probably would not occur in the usual school that the correlations of the 6th and 4th grades would be relatively as high as in this particular school. Assuming that for the usual school there is a progressive gain in correlation with first-year standing as one proceeds from the 4th to the 7th grade, we would have correlations about as follows:[1]

	F_A	7_A	6_A	5_A	4_A
7_A	.67				
6_A	.58	.67			
5_A	.53	.58	.67		
4_A	.50	.53	.58	.67	
σs	σ_{F_A}	σ_{7_A}	σ_{6_A}	σ_{5_A}	σ_{4_A}

The regression equation based upon this table is as follows:

$$F_A = .4616\frac{\sigma_{F_A}}{\sigma_{7_A}}(7_A) + .1458\frac{\sigma_{F_A}}{\sigma_{6_A}}(6_A) + .0910\frac{\sigma_{F_A}}{\sigma_{5_A}}(5_A) + .1094\frac{\sigma_{F_A}}{\sigma_{4_A}}(4_A) \qquad (a)$$

In case the σ's are all equal this equation becomes, to a very close approximation,

$$54.9(F_A) = 25(7_A) + 8(6_A) + 5(5_A) + 6(4_A) \qquad (b)$$

Equation (a) is the equation recommended for use in the ordi-

[1] See Appx., pp. 91–92.

nary school system. The elementary student of statistics can use this equation without difficulty. First calculate the standard deviations, σ_{F_A}, σ_{7_A}, σ_{6_A}, σ_{5_A}, σ_{4_A}, then express $\left(.4616\frac{\sigma_{F_A}}{\sigma_{7_A}}\right)$ as a single number, and do similarly with $\left(.1458\frac{\sigma_{F_A}}{\sigma_{6_A}}\right)$, etc. This will result in an equation of the type (b) except that the coefficient of FA is unity. It then only remains to substitute the values 7A (the average grade expressed as a deviation from the mean), 6A, 5A, etc., for each individual considered, to obtain the probable grade, expressed as a deviation from the mean, of the individual in his high school work. A similar procedure may be followed for each high school subject, substituting for 7A, 6A, 5A, 4A, the corresponding 7th, 6th, 5th, and 4th grade marks in the subject in question. The result thus obtained will give the relative distribution of the pupils, but in this latter case the most probable mark for the first-year grade may be expected to be numerically a little smaller than the grade given by substitution in the equation.

This amounts to saying that the weighting of the grades of the various years of the elementary school is probably the same whether one deals with average grades or with grades of certain subjects, but that the correlation found is probably smaller in the latter case than in the former. The essential problem is to divide the pupils into groups according to ability, and this the values obtained by substitution in the equation will do with considerable accuracy. The exact degree of accuracy can be determined at the end of the school year by calculating the coefficient of correlation between the prophesied grade and the grade actually obtained by the pupil, due allowance being made for difference in the rigidity of grading in the various sections of the same course—such differences undoubtedly being present if the sections have been divided upon the basis of ability.

At first glance the fact that in the equation (b) the record in the 4th grade is weighted more heavily than the record obtained in the 5th grade is surprising. This arises from the fact that the 4th grade record has a greater independence than the 5th and 6th grade records, and therefore contributes more of an independent nature upon which to estimate freshman standing. This is to

say that from the 4th, 6th and 7th grade records a closer estimate of the 5th grade record can be obtained than can be obtained of the 4th grade record from the 5th, 6th and 7th grade records. The relatively greater independence of the first and last terms of the series is to be expected, and is a cause of their greater weighting.

Before leaving this subject, it is interesting to note that the correlation between the average first-year standing and the average marks for the 4th grade is .624. This high correlation, together with the fact that those who skipped grades were graded high by giving them the grades of the preceding year,[1] instead of being graded low by giving them the grade of the following year, on the ground that having missed a year they would be handicapped in all their succeeding work, is strong evidence that natural capacity is a very much more important factor than training in determining relative scholastic standing. Indeed, it seems that an estimate of a pupil's ability to carry high school work when the pupil is in the 4th grade may be nearly as accurate as a judgment given when the pupil is in the 7th grade, for the correlation in the former case is .62 and in the latter only .10 higher.

[1] See Appx., p. 89.

SECTION 4.—TEACHERS' ESTIMATES AND THEIR SIGNIFICANCE

Toward the close of the first half year the teachers in School A were given lists of pupils in groups 1, 2, and 3 and asked to grade the pupils according to intellectual ability (I. a.) on each list 1, 2, 3, etc., as far as valid judgments could be made. Then, beginning with the weakest, pupils were to be graded a, b, c, etc., as far as judgments could be made. Finally, the remainder of the pupils known to the teacher were to be marked M, signifying a medium group. The demand that ranking be from the best to the poorest, without a medium group, would probably have resulted in less accurate judgments throughout the entire series, for it would have been beyond the power of the majority of teachers to have made valid distinction throughout this range. As it was, on the average, about 25 per cent were placed in the medium group. These rankings were then expressed as deviations from a mean and the results of the gradings by the various teachers combined for each pupil into a single measure.[1]

The same procedure was followed for the traits conscientiousness (Cons.), emotional interest in school work (Emo. i.), and oral expression (Exp.).

These estimates were obtained before the time of the English and history courses used in this study and before the second half year of the mathematics courses used. None of the estimates used were from the mathematics instructors of the pupils. A further effort was made to eliminate the possibility of the estimates of the teachers being more highly correlated than a chance selection of teachers' estimates would yield, by excluding the estimates of English and history teachers who later had the same pupils, in courses here utilized, that they had already taught in the first half year's work. This was possible in all the 460 cases, except in the case of 23 estimates which it was necessary to use in order to secure sufficient data.[2] The teachers' estimates are

[1] See Appx., pp. 92–93.

[2] See Appx. p. 93.

therefore practically free from any direct bearing upon mathematics, English and history, but there is a certain amount of direct connnection with average class standing. For example, the estimate of a teacher of Latin, made near the close of the first half year of school, enters into the teachers' estimate grade, and the grade given by this same teacher for the yearly grade in Latin of the pupil enters into the average grade for the year. This lack of entire independence operates to slightly raise the correlation between teachers' estimates and average class standing. From the data at hand this increase is estimated to be less than .03.

The correlation between gradings by different teachers of the same pupil for the same trait are as follows:

$$r_{(\text{I. a. according to one teacher's estimate})} = .28$$
$$r_{(\text{I. a. according to a second teacher's estimate})}$$

$$r_{(\text{Cons. according to one teacher's estimate})} = .38$$
$$r_{(\text{Cons. according to a second teacher's estimate})}$$

$$r_{(\text{Emo. i. according to one teacher's estimate})} = .31$$
$$r_{(\text{Emo. i. according to a second teacher's estimate})}$$

$$r_{(\text{Exp. according to one teacher's estimate})} = .29$$
$$r_{(\text{Exp. according to a second teacher's estimate})}$$

On an average there were about two and one-half estimates per pupil, so that the reliability coefficients of the various gradings are:

Reliability coefficient of I. a. grading = .493
" " " Cons. " = .605
" " " Emo. i. " = .505
" " " Exp. " = .529

The conditions laid down for securing teachers' estimates were simple to use, but at the same time allowed for as detailed judgment as possible. When first-class teachers can estimate intellectual ability with a reliability of only .29 it lessens the confidence that can be placed in such estimates of ability and conclusions drawn from studies depending upon them.

Teachers' estimates of pupils have the unique value of indicating, more or less accurately, single mental traits instead of a complex of traits such as are involved in the securing of a grade in a subject, but it is highly desirable that these estimates be made by several competent individuals, otherwise the measures are very unreliable.

The correlations between the various estimates and the average grade (Av.) of each pupil are given in the accompanying table. The regression equation based upon this table is:

	Av.	I. a.	Cons.	Emo.i.
I. a.	.72			
Cons.	.62	.61		
Emo. i.	.58	.61	.66	
Exp.	.63	.82	.55	.59

$c \cdot \text{Av.} = 8$ I. a., $+4$ Cons., $+2$ Emo. i., $+1$ Exp.[1]

The correlation between average class standing and the regression equation combination of the estimates of traits

$$r_{\text{Av. (I. a., Cons., Emo. i., Exp.)},} = .76.$$

With such a high correlation, a division of pupils into classes by means of teachers' estimates would be highly reliable. The use of the equation to estimate probable class standing for a school system with a different system of marking from that of the schools here considered follows the same general lines as in the case of elementary school grades. (See Appendix.)

In so far as all high school subjects are equally dependent upon the traits, intellectual ability, conscientiousness, interest and expression, classification according to grading in them is not selective. The extent to which these factors have a common importance for different subjects can be measured by calculating the regression equations involving class standing in different subjects and the teachers' estimates. The following equations give the regression of mathematics, English,[2] and history, respectively, upon I. a., Cons., Emo. i. and Exp. For simplicity, all standard deviations are assumed equal. $M_{t.e.}$ stands for the most probable mathematics grading, based upon teachers' estimates. $E_{t.e.}$ and $H_{t.e.}$ have similar meanings.

$M_{t.e.} = .460$ I. a. $+ .114$ Cons. $+ .129$ Emo. i. $- .014$ Exp.

$$r_{M M_{t.e.}} = .61$$

$E_{t.e.} = .336$ I. a. $+ .251$ Cons. $+ .068$ Emo. i. $+ .083$ Exp.

$$r_{E E_{t.e.}} = .64$$

$H_{t.e.} = .450$ I. a. $- .024$ Cons. $+ .305$ Emo. i. $- .287$ Exp.

$$r_{H H_{t.e.}} = .46$$

[1] See Appx., p. 94.

[2] See Appx., pp. 94–95.

The negative significance of expression in the case of mathematics and history is probably entirely due to the lack of independence of the estimates of oral expression. This is the most objective of the four traits and for that reason it might be considered the easiest to estimate. This view seems incorrect, for oral expression is probably a trait which teachers do not think much about and which they make little attempt to measure, with the result that, when called upon to give an estimate of it, they rely upon associated characteristics, such as intellectual ability, conscientiousness and interest, traits already evaluated in their minds. The result of such a procedure is to obtain measures of expression which are correlated to an unwarrantable degree with more fundamental traits. The tendency to rely upon secondary criteria in the estimation of mental traits is a very difficult one to overcome and the intercorrelations between the four traits estimated are probably all higher than would be shown to be the case with more accurate measurement of them.

The effect of unwarrantably large intercorrelations upon the regression equation is to tend to give the factor which is the dependent one, *e. g.*, in this case expression, small or negative weighting. The weighting which the regression equation gives is the best available for the measure, but the measure is probably not at all an accurate one of the trait considered. A reference to the table on page 16 shows surprisingly high correlations between expression and the other traits. This is itself an indication that the measures called "oral expression" are dependent complexes of other more fundamental traits.

The equations show a decided variation in importance of the different traits with reference to different subjects. Intellectual ability is most important in its bearing upon mathematics and least important in its bearing upon English. Conscientiousness is most important in its bearing upon English and least in its bearing upon history. Interest is most important in its bearing upon history and least upon English. Expression is the most important in its bearing upon English and least in its bearing upon history.

The striking importance of interest for history work, of conscientiousness for English, and of native capacity for mathematics are points which can be utilized by the teacher giving the

instruction as well as by the person attempting to diagnose differentially the pupil's capacities.

Such estimates of teachers are not proposed as a good basis for the determination of the idiosyncracies of pupils, although it is possible in a small way to say what study a pupil will be most efficient in, simply upon the basis of teachers' estimates of general capacities. The correlation between the differences in mathematics and English grades and the differences in estimate of the same is as follows:

$$r_{(M-E)(M_{t.e.}-E_{t.e.})} = .12\,^{1} \quad \text{(Probable error } = .05)$$

It is therefore apparent that the practical value of such teachers' estimates as are here used lies, in the main, in their power to measure general ability, rather than in a power to indicate points of individual strength or weakness. They probably would perform much the same function in connection with vocational guidance.

---

[1] Taking into account differences in standard deviations, *e.g.*, by reducing all standard deviations to unity.

SECTION 5.—SPECIAL TESTS AND THEIR SIGNIFICANCE

The data here concern the same groups of pupils as in the preceding sections dealing with teachers' estimates. Three kinds of tests were given to determine ability, interest and preparation. We may say that three main factors enter into the production of a grade: (1) the mental capacity of the individual, (2) the preparation of the individual for the particular course, and (3) the effort and interest of the individual in the particular subject. The importance of these three factors differs materially for different subjects, *e. g.*, it requires a previous preparation in algebra and analytical geometry in order to carry calculus, and it would be a very peculiar genius who could read Virgil without previous Latin study. In these courses the factor of preparation is of prime importance. Courses in English, history, the sciences, commercial branches and the like, do not so definitely demand a certain accumulated store of knowledge as a foundation. The relative importance of these three factors is worthy of extended study and the data here given throw but little light upon the question. In drawing up tests these three factors were considered with rather special emphasis upon the first and third. The high school subjects covered are algebra, geometry, English, and History, and a description of the various tests will reveal the parts devised to measure each one of these factors.

(A_t) Algebra Test

The following test (called an algebra test simply because it was given to classes just starting in algebra) was devised for the purpose of measuring the ability and preparation of the pupil for algebra. Following each problem are given directions for grading it.

Administration of the test: The only precaution that need be indicated in administering the test is that the teacher refrain from explaining any of the questions verbally. Question 5, in particular, loses its value if the slightest explanation is made.

Algebra Test and Directions for Grading

For all problems: maximum grade 10, minimum grade 0

Name..... Date.....

1. Add:
132
580
649
356
774
263
925
191
417
828

2. Multiply:
42976
30851

From 10 deduct 4 for each error in addition, or in carrying forward.

From 10 deduct 4 for each mistake in placing partial products, 2 for each mistake in partial product, and 2 for each mistake in addition.

3. Divide 457219 by 638 and carry answer to one decimal place.

From 10 deduct 2 for each failure to draw down, 3 for a mistake in decimal point, 2 for each error in subtraction, and 2 for failure to carry work to one decimal place.

4. Simplify: $6\frac{3}{5}-3\frac{1}{4}-1\frac{11}{40}+2\frac{1}{4}-1\frac{3}{4}$.

From 10 deduct 2 for each mistake in simplifying a term, 3 for each error in addition.

Give a grade of 5 for answer $\frac{33}{5}-\frac{13}{4}-\frac{61}{40}+\frac{9}{4}-\frac{7}{4}$.

5. Is the square constructed on a line (3+5) eight inches long greater than, equal to, or less than, the sum of the squares constructed on lines 3 inches and 5 inches long? Explain.

If explanation shows an understanding of the problem grade it 10. Give grade of 2 for $(8\times8)-(5\times3)=49$.

6. What is that number such that if it is multiplied by itself and added to 11 the result is 27?

Give grade of 2 for answer 16.

7. A certain balloon without its basket will just lift a weight of 160 pounds. What is the weight of the balloon and basket if the basket weighs 20 pounds?

Grade 10 for answer -140, or for statement "140 pounds less than nothing," or similar statement. Grade all other answers 0.

8. Simplify: $\frac{7}{16}\times\frac{1}{5}\times\frac{2}{3}\div\frac{7}{30}$.

From 10 deduct 6 for incorrect inversion, or failure to make correct inversion, 4 for each error in cancelling, and 4 for each other error.

9. Simplify: $\dfrac{\frac{3}{4}-\frac{3}{5}}{\frac{2}{3}+\frac{3}{8}}$.

Give credit of 3 for correct simplification of numerator, 3 for denominator, and 4 for the balance.

10. Simplify: $\dfrac{\frac{2}{3} \times \frac{11}{12}}{\frac{1}{9} \times 5\frac{1}{2}}$.

Same as 9.

11. Find lowest common multiple and highest common factor of 42, 56, 63, 84.

Give credit of 2 for all factoring correct, 4 for H. C. F. if plainly labelled and 4 for L. C. M.

12. January 1 of a certain year the temperature was 70 degrees, January 2 it was 40 degrees. What was the temperature January 3 if it was still colder and the difference between the temperatures of January 3 and January 1 was three times as great as the difference between the temperatures of January 2 and January 1?

Give credit of 2 for answers of 90°, or 20°.

13. Find a number such that if 5 is added to 3 times that number the result is 38.

Grade 10 or 0.

14. One-third of a certain number added to 7 is equal to 22. What is the number?

Give credit of 2 for answer of 5.

The primary purpose of problems 1, 2, 3, 4, 8, 9, 10, 11 is to test the thoroughness of the preceding preparation of the pupil. The remaining questions are primarily for the purpose of testing his capacity to deal with algebraic material; questions 5, 6, and 12 testing his ability to understand the written statement of a problem and to deal with negative magnitudes, and questions 6, 13, 14 demanding elementary algebra, or at least a process of thinking which is very closely related to simple algebraic reasoning. Problem 7 may be objected to on the ground that the student of physics, who has weighed gases and the like, may be misled by the term "weight" when a negative magnitude is demanded for a correct solution of the problem. None of the subjects showed that this particular difficulty was present in their minds, probably because none of them were familiar with the necessary physics.

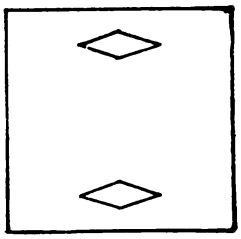
The grading is highly objective and the reliability accordingly very high. To calculate the reliability, a sample was graded by two judges. The correlation between the total grades for each pupil as determined by the two judges, is .996, which is the reliability coefficient, as most of the papers were graded by a single judge. The various problems in this test are approxi-

mately[1] of equal significance and the sum of the grades of all of the problems is the grade for the test. To obtain a distribution which would be convenient for purposes of calculation from the grade thus obtained, the average grade of the group in question was subtracted and the remainder divided by 5, keeping the result to the nearest integer. This grade is designated by A_t (algebra test) and when grouped with the grades of the geometry test by M_t (mathematics test).

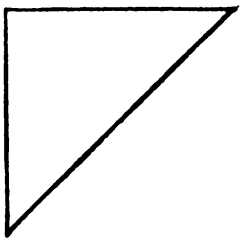
(G_t) Geometry Test

Administration of test: In giving the following test, problems 1 and 2 require explanation. A demonstration is given the pupils of a simpler problem, to show the nature of the requirement.

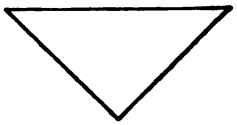
Paper, about one foot square, is held against the blackboard with the top edge horizontal, folded once from the bottom up, a second time from the right to the left and, while still close to the blackboard and without rotating the paper, a V-shaped notch cut into it. The pupils are then asked to describe, orally, the appearance of the paper when unfolded and after receiving a few correct answers the paper, still against the blackboard, is unfolded, enabling the entire class to see that the unfolded paper does have the shape



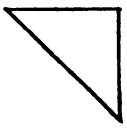
The test question is then given, folding the paper along a diagonal, giving it this appearance



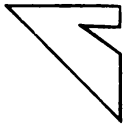
a second fold leaves it in this shape



a third in this



and after notching it appears thus



The pupils are then asked to represent its appearance when unfolded.

The nature of the requirement in the second question can be made clear by means of two large wooden compasses, each hand holding points of the different compasses, and demonstrating the

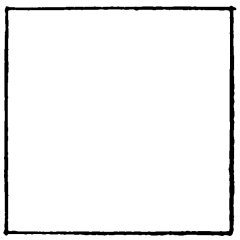
[1] See Appx., pp. 95–96.

lack of rigidity of the diamond-shaped frame thus formed. Care should be taken not to indicate the position, or number, of braces necessary to make the figure rigid.

Geometry Test and Directions for Grading

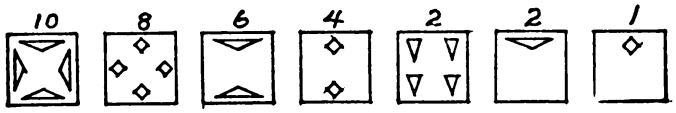
All problems: Maximum grade 10, minimum grade 0

Name..... Date.....



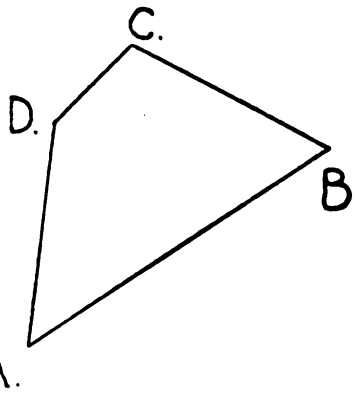
1. The accompanying diagram represents a square sheet of paper which is folded three times by the teacher and cut. Draw in the square, in their correct position, the holes cut out.

Credit given for drawings as follows:



2. Suppose that *AB*, *BC*, *CD* and *DA* are sticks of wood hinged together at points *A*, *B*, *C* and *D*. How many braces are needed to make the figure rigid and where would you put the brace or braces?

For one brace properly placed give credit of 10, improperly placed credit of 4. For two braces give credit of 2.

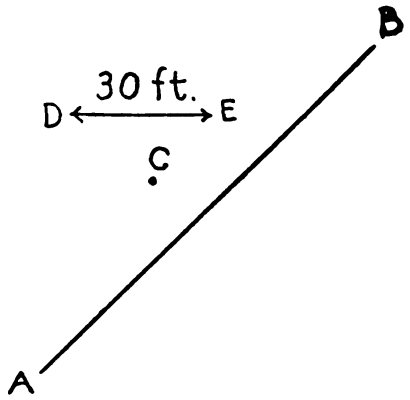


3. John's mother forbids him to leave Broadway. James' mother forbids him to leave Amsterdam avenue. They are obedient sons, but are also very fond of seeing each other, so how can they meet?

Give credit 10 for answer 72nd St. Any other answer 0.

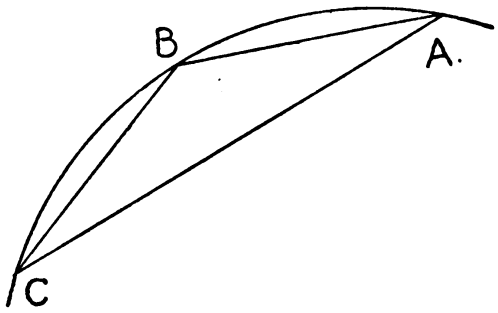
4. *AB* is a railroad track. *C* is a stake to which a cow is tied with a 30 ft. rope. *DE* measures a distance of thirty feet. One day the cow, while grazing at the end of its rope, is struck by the train. Mark the place or places in the diagram at which this must have occurred

Give credit 10 for indicating two points correctly, 6 one point, 6 for distance between the two correct points.

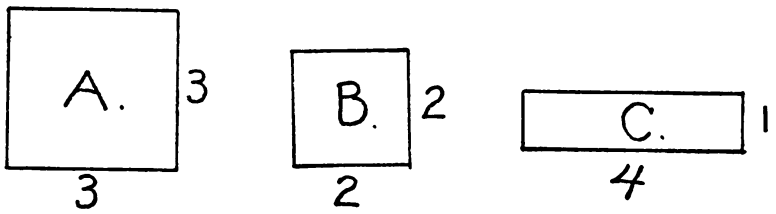


5. In the accompanying figure, the distance along the circle from B to C equals the distance along the circle from A to B. Also, the straight line BC equals the straight line AB. Does the curved line from A to C equal twice the curved line from A to B? ..... Does the straight line AC equal twice the straight line AB? .....

Give credit 10 for both answers correct, 2 for one correct.



6. What are the areas of figures A, B, and C? A....., B....., C.....
In what respect, if any, are figures A and B alike? .....
In what respect, if any, are figures B and C alike? .....
In what respect, if any, are figures A and C alike? .....



Give credit:
1 for $A=9$, $B=4$, $C=4$.
3 A and B both squares.
2 A and B alike in shape.
3 B and C equal in area.
1 B and C both parallelograms.
3 A and C both rectangles.
2 A and C not at all alike.

7. Find the value of x from the equation $\frac{x}{b}=\frac{c}{d}$.

Find value of y from the equation $\frac{a}{y}=\frac{c}{d}$.

Give credit:

4 for $x=\frac{bc}{d}$.

6 for $y=\frac{ad}{c}$.

Anything else 0.

In the following question, is the third statement proved if the first two are true? Write "proved" or "not proved" after it.

8. 1. Every chalet is a bungalow.
2. Jones' house is a bungalow.
3. Therefore Jones' house is a chalet. .....

Credit 10 or 0.

Do the same for the following question:

9. 1. Every chalet is a bungalow and every building that is not a bungalow is not a chalet.
   2. Smith's house is a bungalow.
   3. Therefore Smith's house is a chalet. .....

Credit 10 or 0.

Do the same for the following question:

10. 1. Every building that is not a chalet is not a bungalow.
    2. Brown's house is in no particular different from a bungalow.
    3. Therefore Brown's house is a chalet.

Credit 10 or 0. .....

The primary purpose of all the problems, except problem 7, is to test capacity; problem 1 testing ability to image geometrical forms and movement. Problems 2, 3, 4, 5, 6, are locus problems, and problems demanding a common sense interpretation of geometric facts. Problem 7 is a problem to test the pupil's previous preparation in the line of ratio and proportion. Problems 8, 9 and 10 are problems in logic aimed to test the pupil's ability to handle reductio ad absurdum proofs and converse propositions. Problem 10 is a difficult problem and is meant to tax the most able pupil. It may be said that all the tests were devised with a view to securing a good distribution of marks. It was desired that the most efficient pupil would just succeed, or just not succeed, in making a perfect score, while at the same time the tests were meant to be sufficiently easy in places to secure the coöperation of the poorest pupil. The reliability coefficient for this test is very high, being .994. A number of the problems in this test did not prove as significant as conferences with various teachers of geometry led the author to expect,[1] and the grade for the entire test is taken as the sum of the marks for problems 1, 7, 8, 9, 10. To obtain a convenient distribution for purposes of calculation, from the grade thus obtained the average of the group in question was subtracted, and the remainder divided by 3, keeping the answer to the nearest integer. This grade is designated by G_t, or M_t when grouped with algebra data.

(E_t) English Test

The English test which follows, together with the explanation of the grading and the sample grading, will explain the purpose for which it was devised. It is fundamentally an ability test and seems to meet the requirements very well, as none of the

[1] See Appx., p. 97.

pupils gave evidence of familiarity with the subject matter of the test.

Administration of the test: Tell the pupils that you are about to read an account of an incident in the life of the founder of one of the great Eastern religions and that after reading you are going to ask them questions about it. Then read the following:[1]

English Test

"A woman—dove-eyed, young, with tearful face
And lifted hands—saluted, bending low:
"Lord! thou art he," she said, "who yesterday
Had pity on me in the fig-grove here,
Where I lived lone and reared my child; but he
Straying amid the blossoms found a snake,
Which twined about his wrist, whilst he did laugh
And tease the quick forked tongue and open mouth
Of that cold playmate. But, alas! ere long
He turned so pale and still, I could not think
Why he should cease to play, and let my breast
Fall from his lips. And one said, 'He is sick
Of poison'; and another, 'He will die.'
But I, who could not lose my precious boy,
Prayed of them physic, which might bring the light
Back to his eyes; it was so very small
That kiss-mark of the serpent, and I think
It could not hate him, gracious as he was,
Nor hurt him in his sport. And some one said,
'There is a holy man upon the hill—
Lo! now he passeth in the yellow robe—
Ask of the Rishi if there be a cure
For that which ails thy son.' Whereon I came
Trembling to thee, whose brow is like a god's,
And wept and drew the face cloth from my babe,
Praying thee tell what simples might be good.
And thou, great sir! didst spurn me not, but gaze
With gentle eyes and touch with patient hand;
Then draw the face-cloth back, saying to me,
'Yea! little sister, there is that might heal
Thee first, and him, if thou couldst fetch the thing;
For they who seek physicians bring to them
What is ordained. Therefore, I pray thee, find
Black mustard-seed, a tola; only mark
Thou take it not from any hand or house
Where father, mother, child, or slave hath died;
It shall be well if thou canst find such seed.'
Thus didst thou speak, my Lord!"
The Master smiled
Exceeding tenderly. "Yea! I spake thus,
Dear Kisagôtami! But didst thou find
The seed?"
"I went, Lord, clasping to my breast
The babe, grown colder, asking at each hut—
Here in the jungle and towards the town—
'I pray you, give me mustard, of your grace,
A tola—black:' and each who had it gave,

[1] Taken verbatim from Edwin Arnold, Light of Asia, p. 124–8.

For all the poor are piteous to the poor;
But when I asked, 'In my friend's household here
Hath any peradventure ever died—
Husband or wife, or child, or slave?' they said:
'O Sister! what is this you ask? the dead
Are very many, and the living few!'
So with sad thanks I gave the mustard back,
And prayed of others; but the others said,
'Here is the seed, but we have lost our slave!'
'Here is the seed, but our good man is dead!'
'Here is some seed, but he that sowed it died
Between the rain-time and the harvesting!'
Ah, sir! I could not find a single house
Where there was mustard seed and none had died!
Therefore I left my child—who would not suck
Nor smile—beneath the wild-vines by the stream,
To seek thy face and kiss thy feet, and pray
Where I might find this seed and find no death,
If now, indeed, my baby be not dead,
As I do fear, and as they said to me."

"My sister! thou hast found," the Master said,
"Searching for what none finds—that bitter balm
I had to give thee. He thou lovedst slept
Dead on thy bosom yesterday: to-day
Thou know'st the whole wide world weeps with thy woe:
The grief which all hearts share grows less for one.
Lo! I would pour my blood if I could stay
Thy tears and win the secret of that curse
Which makes sweet love our anguish, and which drives
O'er flowers and pastures to the sacrifice—
As these dumb beasts are driven—men their lords.
I seek that secret: bury thou thy child!"

The questions to the pupils are as follows:

(1) State one or two things about the story which you particularly liked. (6 minutes.)

(2) Write an account of the story as fully as you can remember it. (9 minutes.) (5 minutes for reading the passage, giving a total time of 20 minutes.)

The grading was upon the following four points, though the grading upon E_v, E_a and E_w only, are used in obtaining a single measure of the English test:

E_v Valuation of the essential ideas in the poem. (Grade approximately from 0 to 10, with an average of 5. Further explanation follows.)

E_a Accuracy and extent of description. (Start with 4 and to this add ½ for each point correctly made and subtract 1 for each point incorrectly made.)

E_w Written expression. (Grade from 0 to 10, with an average of 5. Give some slight weight to spelling.)

E_d Dramatization. (Grade from 0 to 10, with an average of 5.)

The reliability coefficient of the English test, $\frac{E_a+E_v+W}{3}$, or E_t, equals .969, since the grades used are the sum of the grades of two judges and the correlation, based on a sample of 36, between the grades given by the judges is .940.

The grading for valuation of the essential ideas of the poem was largely based upon the answer to question (1). It followed closely the following scheme:

Grade below: given for selecting the following as the point liked the best:

10±1 Love shown in the Master's way of teaching that suffering is universal.
8 ±1 "The grief which all hearts share grows less for one."
8 ±2 Appreciation of poetry and language used.
6 ±1 Master's statement to the mother that the whole world suffers, or the idea that suffering is universal.
5 ±1 Master's statement, "I would pour my blood if I could stay thy tears."
4½±1 "The poor are piteous to the poor."
4 ±1 Mother's tenderness and love for her child, or the mother's cry that she cannot lose her child.
3 ±1 Mother's trust in the Master, and her courage.
1 ±1 Honesty and strength of mother in not taking forbidden seed.

The plus or minus after each grade indicates the amount that quite generally is to be added or subtracted, depending upon the answer to the second question, as follows:

Add 1 for genuine appreciation of the Master's character, *i. e.*, his gentleness, compassion, humanity, and for genuine appreciation of the first three points above.

Add 0 for correct narrative.

Add −1 for demonstrated lack of correct appreciation and for incorrect narrative which betrays incorrect appreciation of the Master's character and motive.

Having grades for accuracy (E_a), valuation (E_v) and written expression (W) $\left(W=\frac{E_w+H_w}{2}\right.$, or the average of the grades for written expression for the English and history tests$\left.\right)$ the single grade for the entire test is the average of these three grades minus the mean for greater convenience after multiplication by two, *i. e.*, $E_t=\frac{2}{3}(E_a+E_v+W-\text{mean})$.[1]

In the following samples, given to illustrate the method of grading, the points for which credit in accuracy has been given are underscored once and the incorrect points, for which credit was deducted, have been underscored twice. The grading for dramatization is not used in the final score for the test, but is given as it may have some interest in itself and probably has a specific significance in some other bearing than upon English.

[1] See Appx., p. 98.

Name (Pupil No. 187). Date.....

	Grading			
	E_a	E_v	E_w	E_d
I. I like the story of the poem very much. I think it	4			
is very pathetic and has a good point—I refer to the				
mothers grief being lessened because she knew all others				
had to suffer. I also like the way the author wrote the				
poem. It is well told, and in beautiful language.		8		
II. The mother throws herself at the feet of the				
master and says that she come to him to now because				
yesterday he was so kind to her and when on the	.5			
desert her child was playing with a snake and it bit	−1			
him, and she feared for his life, she went to him and asked	.5			
consul of him. He said that she should get mustard seed	.5			
from somebody, but beware not to take it from anyone				
in whose house there has been a death.	.5			
"Yes," says the master, "and did you get the mustard				
seed?"	.5			
"I went into every hut," says she, and	.5			
asked for the mustard seed and the good people	.5			
all gave it to me, for the poor give to the	.5			
poor, but when I asked if there had been a death	.5			
in the house, they all answered, yes. So now I have	.5			
to you, dear master, to once more seek consul, ere				
my baby dies, if he be not dead already."	.5			
The consul speaks, "Yesterday when you came to				
me, and asked my aid, your babe was already dead,	.5			
but I let you go and try to get the mustard seed, so		1		
	9	9	4	7
H_w			4	
$\frac{E_w+H_w}{2}$			4	

$E_t = \frac{2}{3}(E_a+E_v+W-\text{mean}^1) = \frac{2}{3}(9+9+4-19) = 2.0.$

[1] For table of means see p. 68.

Name (Pupil No. 188). Date.....

	GRADING			
	E_a 4	E_v	E_w	E_d
I. I think that the lord was very good and tender in telling the mother to find something that was not, and telling her that that thing would cure the baby. In that way she did not get a shock.	.5	10		
I think the mother did right by not giving the baby the wrong seed, when she could not find the right seed and then going to the lord again for advice.				
II. When the story opens; a young mother is bending low to a lord of her religeon and tells the following story: "Yesterday while my child was playing in the garden and finding a snake began to play with it.	.5			
	.5			
The snake bit it leaving a small mark on the child's	.5 .5			
skin and some neighbors said that the baby was dead,	.5			
others that it was not dead yet but was dying of poison.	.5			
And others advised me to go to a certain lord who lived over the hill and to ask him what to do to the baby to save him and then they told me that the man in a yellow robe who was walking on the road by the house was the lord so I ran up to him and knelt down to salute him and then I took the face cover off his face	.5			
	.5			
and showed the lord the mark and the lord felt the baby	.5			
and told me to go and get some black mustard seed	.5			
but it must not come from a house wherein a father, mother, child or	.5			
	10.5	10	2	7
H_w			4	
W			6	

$E_t = \frac{2}{3}(10.5+10+3-19) = 3.0.$

Name (Pupil No. 200). Date.....

	Grading			
	E_a	E_v	E_w	E_d
	4			
I. It seems to me that one of the finest points in this passage, is the tenderness that runs all through it. It is in the mother's speach when she talks of her child, and also in the answer,—full of pity and truth. Another thing about it is the wording, for the passage, although written in prose, sounds almost like poetry or music, on account of the beautiful words used.		9		
II. A woman, with tears in her eyes, came and	.5			
threw herself at the feet of the Master. She told	.5			
him of her child, who, while playing in the garden				
had been stung by a serpent. In a short time he	.5			
grew pale and cold, and people said he would surely	.5			
die. But the mother, unwilling to give up hope, had	.5			
gone to the Master, certain that he, in his wisdom, could aid her, and restore her boy to health. And				
the Master, full of compassion, had told her to go	.5			
with her child, from house to house, and beg for	.5			
mustard seed, but from any house where man or woman, slave or child had died, she should not accept				
the seed, but go elsewhere. Carrying her child the	.5			
mother had begged at every door, but, although	.5			
glad to give her seed, in every house someone had	.5			
died, so the		1		
	9	10	6	6
H_w			9	
W			7.5	

$E_d = \frac{2}{3}(9+10+7.5-19) = 5.0.$

Name (Pupil No. 206). Date.....

	Grading			
	E_a	E_v	E_w	E_d
	4			
I. The story is extremely touching. The picture of the the "sad tearful face" is most vivid, and the little	.5			
"face so pale and still" can almost be seen. We	.5			
certainly hope that the mother was comforted.		6		
II. A woman once had a beautiful little child. She loved him dearly. The boy would daily sleep on some				
dry grass. One day, the mother being tired and weary,	−1			
fell asleep. (How) Then came a large venomous	−1			
snake and bit the child, so that he died. And when	−1			
the mother awoke she and saw the still pale face of her	.5			
baby she rent her garments, whilst all the people said	−1			
"There is no hope." But then one man espied the doctor of the village. And quickly he spoke, saying "There is the doctor! go you to him, mayhap he'll give your child to you again." Then on her knees,				
saluting humbly she begged the "wise" man. "My	.5			
daughter, if you wouldst your child recover go you from house to house and beg and say "Give me some				
mustard seed" but if there be by any chance a dead	.5			
man in that house, then leave the seed, for is a child	.5			
or man or woman there has died, the charm is broken. Go, and if thou succeedest—				
	3.0	6	5	10
			6	
H_w				
W			5.5	

$E_t = \frac{1}{2}(3+6+5.5-19) = -3.0$

Name (Pupil No. 124). Date.....

	Grading			
	E_a	E_v	E_w	E_d
I. After the snake bit the child how very sad the mother was, and she did every thing to try and save him.	4 .5	2		
II. The story was about this child that was bitten by a snake and he was going to die and the mother tried to get some mustard seed but nobody had any and the child died.	−1 .5 −1	−1		
	3	1	0 0	3
H_w				
W			0	

$E_t = \frac{2}{3}(3+1+0-15) = -7.3$

(H_t) History Test

The history test follows much the same lines as the English test. None of the pupils showed evidence of familiarity with the subject matter.

Administration of the test: Draw a map of Italy and Sicily on the blackboard, indicating the following provinces and cities: Piedmont, Genoa, Venice, Rome, Naples, Calabria, Messina, Palermo and Marsala. In addition to these places write the names "Garibaldi" and "Victor Emmanuel" on the board. Give orally necessary historical groundwork as follows: "Victor Emmanuel was the king of Piedmont. (Point out.) He had expressed his willingness to lead an insurrection to establish a free and united Italy whenever the other states of Italy should revolt. The Two Sicilies (point out) and the other small states of Italy were governed by rulers who were opposed to a united and republican Italy.

"These are all the facts that you will need to know to understand the selection about to be read. Remember those points that you like and that you consider historically important." (Allow 8 minutes for the reading.)[1]

History Test

"Garibaldi was the hero on the field of battle. The last of knight-errants, he was the very incarnation of Romance and Revolution. Bred to the sea,

[1] This selection is a modification of pp. 392–3, 402–4, Sedgwick, A Short History of Italy.

he always retained the jaunty, gallant bearing of a mariner. His countenance (childlike and lionlike)—with its broad tranquil brow, benign eye and resolute mouth—in youth all sparkling, gradually changed with care and disillusion, but he still kept the seaman's mien and the seaman's lightsome eye. He was the beau ideal of a romantic hero. After his unsuccessful raid into Piedmont he had gone to South America, where he lived a wild life of guerilla warfare, fighting like a Paladin on behalf of republican revolutionaries who were struggling for their freedom. All the time he was training a band of Italian adventurers, his legion, so that they should be ready when their country had need of them. These men rushed to the defense of the city. Their entry was most picturesque. The gaunt soldiers, wearing red shirts and pointed hats topped with plumes, their legs bare, their beards full-grown, their faces tanned to copper color, with their long black hair dangling unkempt, looked like so many Fra Diavolos. At their head Garibaldi, in his red shirt, with loose kerchief knotted round his throat, the regular beauty of his noble, leonine face set off by his waving hair, mounted on a milk white horse, rode like a demigod."

"A short time after what has just been read the following events took place:"

"In the meantime Francis II, a weak, ignorant, bigoted lad, had mounted to the throne of the 'Two Sicilies.' In April, 1860, a revolt began in Palermo, and, though suppressed there, spread. Two young patriots, Crispi and Pilo, went about stirring the people to action. Garibaldi was begged to put himself at the head of the proposed revolution. On the night of May 6, two ships, the *Lombardy* and the *Piedmont*, secretly left Genoa, and took Garibaldi and a thousand volunteers aboard. This band, known as 'the thousand,' is nearly as famous and as legendary as King Arthur and his Round Table. On May 11, the ships landed at Marsala. Two cruisers from Naples came up, but two English men-of-war happened to be there also; and the English captains, under guise of friendly notifications to the Neapolitans, took some action which delayed the latter long enough to let the last Garibaldians disembark. Once on shore Garibaldi's volunteers ran to secure the telegraph office. They arrived just after the operator had telegraphed that two Piedmontese ships, filled with troops, had come into the harbor; a Garibaldian was able to add to the message, 'I have made a mistake; they are two merchantmen.' The answer came back, 'Idiot.' The volunteers marched inland. A provisional government was organized; Garibaldi was made dictator, and Crispi secretary of state. The cry was 'Italy and Victor Emmanuel!' Garibaldi was joined by insurgent Sicilians, and, with numbers considerably increased, fought and defeated the Bourbon army. The story reads like the exploits of Hector before the Greek trenches. Victory followed victory. Palermo fell, Milazzo and Messina; then he crossed the straights and invaded Calabria. This marvelous triumph, for there had been thirty thousand troops to oppose Garibaldi, frightened King Francis; he proclaimed a constitution, but it was too late. Garibaldi swept on victorious, and the king fled from Naples (Sept. 6); the next day Garibaldi marched in and assumed dictatorship of the kingdom.

"Victor Emmanuel took up the cause and, marching south, joined with the forces of Garibaldi and together they decisively defeated the opposing army. In February, 1861, the first Italian parliament was held and Victor Emmanuel formally received the title of 'King of Italy.' Excepting Rome and Venice, Italy was free and independent."

Questions as follows:

(1) What do you think was probably the next important event? (2 minutes.)

(2) Describe the character and appearance of Garibaldi. (4 minutes.)

(3) Beginning with the departure of Garibaldi and his men from Genoa write a detailed account of as much of the story as you have time for. (Balance of time—6 minutes.)

Each paper was graded upon the following points, though the grading upon H_a, only, is used as the measure of the history test. H_w is used in connection with the English test.

H_v Valuation—historical forecast and appreciation of the essential historical facts. (Grade approximately from 0 to 10, with an average of 5. Further explanation follows.)

H_a Accuracy and extent of description. (Start with 4 and to this add ½ for each point correctly made and subtract 1 for each point incorrectly made.)

H_w Written expression. (Grade from 0 to 10, with an average of 5. Give some slight weight to spelling.)

H_d Dramatization. (Grade from 0 to 10, with an average of 5.)

The grading for valuation depended in part upon the answer to question (1). The grading of this historical forecast followed closely the scheme below:

Grade below: given for selecting the following as the next historical event:

7 ± 1	Drawing up a constitution.
6 ± 1	Peaceful acquisition of Rome and Venice.
5 ± 1	Conquering of Rome, or of Rome and Venice. Garibaldi given some honor. Establishment of a government.
4 ± 1	Peace for a short time and then revolts. Garibaldi rebels against Victor Emmanuel.
3 ± 1	Failure to answer.
2 ± 1	Uprising and revolution by the people. Garibaldi made king.

The plus or minus after each grade indicates the amount that quite generally is to be added or subtracted, depending upon the answers to the second and third questions. For the second question,

Add 1, or more, for genuine appreciation of the traits of character which were essential to Garibaldi's success.

Add 0 for correct personal description.

Add −1 for incorrect personal description which does injustice to Garibaldi's character, and for irrelevant but correct data, e. g., trip to South America.

For the third question,

Add 1 for correct references to motive and organization, e. g., "to establish a free and united Italy"; "formed a provisional government"; "joined by insurgent Sicilians"; "proclaimed a constitution"; etc.

Add 0 for correct narrative.

Add −1 for incorrect narrative which violates principles involved and for misunderstanding of organization, e. g., "Garibaldi sailed to Marseilles and fought the King of France"; or for attributing incorrect motives.

As already mentioned the accuracy grading H_a is the measure for the entire test for convenience multiplied by two, i. e., $H_t = 2(H_a - \text{mean})$.[1]

[1] See Appx., p. 99.

The reliability coefficient of H_t is .956 since the grades used are the sum of the grades given by two judges and the correlation between the latter, based on a sample of 36, is .916.

In the following sample tests, given to illustrate the method of marking, the gradings for valuation and dramatization are included though they are not used in obtaining the final score for the test.

Name (Pupil No. 226). Date.....

	Grading			
	H_a	H_v	H_w	H_d
I. The next important event might have been the	4			
taking of Venice and Rome by Garibaldi and Victor Emmanuel, making Italy a United Kingdom.		5		
II. Garibaldi wore a red coat, high hat with a plume	.5			
and boots. His character was a very strong and good	−1			
one. He desired to do good for his country, therefore	.5			
fought well and won many victories.	.5	.5		
III. Garibaldi was a very brave soldier who was				
sent by the King of Piedmont, Victor Emmanuel. He	−1			
and one thousand men started down to the southern	.5			
part of Italy, by water. They reached the most south-				
ern of the two Sicilys, and conquered Palermo first.	−1			
A telegraph had been sent just before their arrival				
and it said that these men were coming to conquer	−1			
them, but as Garibaldi reached there he said "no that				
is not right, we are merchants," so another telegram	−.5			
was sent. Garibaldi and his men marched on to				
Marsala and conquered that city. He won victory . . .		−.5		
	1.5	5	5	5

$H_t = 2(H_a - \text{mean}) = 2(1.5 - 6.5) = -10.0.$

Name (Pupil No. 225). Date.....

	H_a	H_v	H_w	H_d
I. The fall of Rome or Venice which would make all Italy independent!	4	5		
II. Garibaldi was a seamanlike, strong phisycaly and mentaly with a patriotic spirit. He had a handsome face, intelligent eyes. Dressed in plumed cap, red shirt and loos bandana around his neck he looked quite fierce.	.5 .5 .5 1.5	0		
III. It took us only a day to coast down to Palermo where we disembarked as quickly as possible while to English cruisers held up some of the ships of the King of Sicily. We captured the town and assured the King of Sicily over telegraph that no such thing had happened. We had soon captured all the southern cities and were marching north to Capture Naples but the King hearing of our approach fled. King Emanuel came to our help and we soon had all of the states except Rome and Venice which we soon expect to have. Our brave leader Garibaldi is Dictator of Sicily.	.5 .5 −.5 .5 1 .5 .5	1		
	10.0	6	5.5	10

$H_t = 2(10 - 6.5) = 7.0$.

Name (Pupil No. 212). Date.....

	Grading			
	H_a	H_v	H_w	H_d
I. The completion of the Kingdom by capturing Rome (especially) and Venice; and making Emanuel's claim to the throne secure.	4	7		
II. Garibaldi, having been a sailor inherited their jovial demeanor and carriage. He was exceedingly adventurous, and, we hear of him in such countries as were engaged in war. He was tanned by exposure as a seaman, strong, muscular and athletic. He inherited a trait of being able to rule and govern men as is shown by his checkered career. He was perpetually clad in a red shirt and sometimes a kerchief of the same brilliant hue, encircled his neck. His hair was luxurious and abundant which set off his appearance nicely.	.5 .5 .5 .5 .5 −.5 .5 −.5 .5 .5 .5	1.5		
III. In April 1860 Garibaldi left Genoa in command of 1000 volunteers embarked on two Piedmontese vessels. He reach Marsala and marched his men landward to Palermo which he captured. Palermo had previously revolted and his campaign was in accordance to it. From Palermo he sailed to Calabria which he captured. He proceeded northwards and was joined by Emanuel. Together thay captured all of Italy except Rome and Venice. Garibaldi assumed the title of Dictator. Francis, King of the two Sicilies abdicated. Carassi was made secretary. Emanuel was proclaimed King and the first Parliament met 1861, which represented all the individual states combined. (Deduct 1 for order of events)	.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 −1	3.		
	12	11.5	9	5

$H_t = 2(12 - 6.5) = 11.$

Name (Pupil No. 234). Date.....

	H_a	H_v	H_w	H_d
I. The fall of Rome was probably the next important event, maybe Venice at the same time.	4	6		
II. Garibaldi was a strong man. As he had been in South America he was toughened to wild life. He was a man who was strong as a general and could command troops well.	.5			
	.5	1		
He had black wavy hair and wore a red shirt.	.5 .5			
III. When Garibaldi left Genoa secretly he took with him a thousand men. These men were like King Arthur and his Knights of the Round Table. When they came near Marsala two war vessels from Naples followed them but at the same time two English men-of-war were there and they delayed the men-of-war from Naples. Garibaldi reached Marsala and went to a telegraph office and was just too late to stop a message saying that they had arrived. He had the operator send the message that he mistook the vessels and that they were only merchant vessels.	.5			
	.5			
	.5			
	.5			
	.5			
	.5			
	.5			
	.5			
	10.5	7	8	3.5

$H_t = 2(10.5 - 6.5) = 8.0$.

Name (Pupil No. 227). Date.....

	Grading			
	H_a	H_v	H_w	H_a
	4			
I. The two Sislies were united, no doubt.		2		
II. Garibaldi must have been stern and had good discipline to manage so many people, and train Italian soldiers. (He was perfectly honest as he said he would help Italy and did.) Garibaldi, was tall,	−1 −1			
had dark wavy hair that fell in waves over his	.5			
crisp black eyebrows, his eyes stood out like black	−1			
diamonds in a white velvet background. Garibaldi	−1			
wore a flaring red shirt.	.5			
III. Garibaldi left Genoa with a crew of a thousand men, sailed directly southward to Palermo; when the	.5 −1			
people there heard that vessels had landed they sent a wireless of the news and also for help. Garibaldi	−1 −1			
arrived just in time to add a few words to the message.	.5			
After he had conquored and captured Messina, and	.5			
the other cities he worked his way north where another force joined him and together they captured Naples.	−1 .5	−1		
	−1	1	7.5	8.5

$H_t = 2(-1-6.5) = -15.$

(M_i, E_i, H_i,) Interest Tests

In attempting to test a pupil's interest in some given high school subject an error would likely be introduced if that subject alone were dealt with, as the pupil would readily see the object of the test and, possibly unconsciously, be influenced thereby. For purposes of tapping a pupil's personal preference it is very much more significant for him to say that of all vocations he prefers that of teaching mathematics, than for him to reply in the affirmative to the question "Would you rather teach mathematics for a vocation than to do anything else?" This illustration serves to emphasize the difference in the nature of response of an individual when he is indicating a spontaneous preference from his response when he is accepting or rejecting a controlled choice.

In order to insure such spontaneity and freedom of choice, an interest test was so devised as to cover impartially all the ordi-

nary interests of a pupil. Because this test covers so broad a field it may be used equally well to measure a pupil's interest in lines other than mathematics, English and history, which are the lines for which its significance has been evaluated in this study. When graded along the line of English, the grade of this test is designated by E_i, along the line of mathematics by M_i, and along the line of history by H_i. The grading of this test is accomplished by means of tables given on succeeding pages.

Administration of test: The pupils were told to answer the questions on the sheets handed them. As much time as was needed was given—most of the pupils finishing the task in 40 minutes.

Interest Tests

Name..... Date.....

1. Go through the accompanying list of magazines and put an x opposite those with which you are not familiar, that is, opposite those of which you have never looked through at least two numbers.

1. All Story
2. American Boy
3. American
4. Argosy
5. Atlantic Monthly
6. Black Cat
7. Blue Book
8. Bookman
9. Cassier's
10. Century
11. Collier's Weekly
12. Commoner
13. Cosmopolitan
14. Country Life in America
15. Craftsman
16. Current Literature
17. Delineator
18. Electrical World
19. Etude
20. Everybody's
21. Good Housekeeping
22. Graphic
23. Green Book
24. Hampton's
25. Harper's Weekly
26. Harper's Monthly
27. Hearst's
28. Home Needlework
29. Ill'd London News
30. L'illustration
31. Illustrirte Zeitung
32. Industrial Eng'g
33. International Studio
34. Ladies Home Journal
35. LaFollette's
36. Leslie's Weekly
37. Life
38. Lippincott's
39. Literary Digest
40. McClure's
41. Metropolitan
42. Modern Priscilla
43. Moving Picture World
44. Munsey's
45. Musician
46. National Geographic
47. Outing
48. Outlook
49. Photographic Times
50. Pictorial Review
51. Popular Mechanics
52. Popular Science Monthly
53. Printer's Ink
54. Puck
55. Red Book
56. Review of Reviews
57. St. Nicholas
58. Saturday Evening Post
59. Science
60. Scientific American
61. Scribner's
62. Smart Set
63. Strand
64. System
65. Technical World
66. Theatre
67. Woman's Home Companion
68. Wilshire's
69. World's Work
70. Youth's Companion

Go through the list again, marking the five that interest you most A, B, C, D and E; A for the most interesting of all, B for the next most interesting, and so on. Do not spend much time in deciding *exactly* upon your preferences.

2. Briefly tell why you particularly enjoy reading the magazine that you have marked A.

3. Name three books which you have read in the last two years that have interested you very much. 1. ..... 2. ..... 3. .....

4. Suppose that you have an hour's leisure time, in what outdoor amusement would you prefer to spend it? .....
5. Suppose that you have an hour's leisure time, in what indoor amusement would you prefer to spend it? .....
6. Of the two amusements named in your answers to questions 4 and 5 which do you prefer? .....
7. If you had the opportunity, which one of the following would you attend, supposing each of them to be first class of its kind? Mark it A.

1. Moving picture entertainment.
2. Circus
3. Football game
4. Baseball game
5. Track meet
6. Musical comedy
7. Vaudeville performance
8. Grand opera
9. Boxing contest
10. Band concert
11. Political rally
12. Light opera
13. Drama
14. Lecture, or stereopticon lecture, on a subject that interests you.

8. What occupation would you prefer as a life work? ..... Which would you like next best? .....
9. In the following list of words mark with a 3 those you know the meaning of perfectly and could define as a dictionary does.

If you can explain in a general way the meaning of the word and would understand it when used in a sentence mark it with a 2.

If you cannot explain its meaning but are vaguely familiar with it, mark it with a 1.

If the word is entirely new to you and unknown, mark it with a 0.

In doing this, go through the list four times, the first time marking the 3's, the second time the 2's, the third time the 1's, and the last time the 0's.

1. simile
2. primary election
3. Mason and Dixon's line
4. creed
5. Acropolis
6. rip saw
7. hydrogen
8. compound interest
9. cube root
10. paradox
11. Saracens
12. I. W. W.
13. Whigs
14. theosophy
15. toga
16. block plane
17. NaCl
18. fissure
19. equation
20. guillotine
21. prose
22. syndicalism
23. H_2O
24. transubstantiation
25. gladiator
26. debit
27. gravity cell
28. strata
29. improper fraction
30. lever
31. ragtime
32. physical valuation of railroads.
33. score (in music)
34. commercial fertilizer
35. Magna Charta
36. voucher
37. ohm
38. string halt
39. fourth dimension
40. piston rod
41. Pythagorean proposition
42. single tax
43. stamen
44. hemstitch
45. Spanish Armada
46. statute of limitations
47. coherer
48. vertebrate
49. parallelogram
50. omelette
51. Reichstag
52. Commerce Court
53. states' rights
54. space bar
55. giblets
56. Australian ballot
57. mollusk
58. perspective
59. fireless cooker
60. mortgagee
61. referendum
62. Formosa

10. Tell what each of the following words means as well as you can.

a. simile
b. cube root
c. improper fraction
d. ragtime
e. physical valuation of railroads
f. commercial fertilizer
g. ohm
h. Pythagorean proposition
i. single tax
j. hemstitch
k. vertebrate
l. parallelogram
m. omelette

The last two questions are not solely for the purpose of testing the pupil's interest, as they test his range of information as well. They constitute a vocabulary test in which the words were chosen because of their specific bearing upon all the usual high school courses and in addition upon religion and politics, in order to cover the intellectual field.

Anyone who has read the answers of a few pupils to the questions in this test must feel that strong individual differences are shown. There is much to indicate that the data are highly significant as evidence of interest, but the problem of expressing this in numerical terms and with reference to specific high school courses is far more elusive than simply the determination that the data are significant. For purposes of evaluation, it would be possible, theoretically, to have a large number of judges grade each pupil's paper upon the various questions with reference to the significance of each of the questions in turn as evidence of interest, severally, in mathematics, English and history. Practically this method is valueless, as judges with sufficient leisure and patience could not be found and because their work would apply only to the actual papers graded and would be of no aid to another party desiring to give the test. A result embodying all the advantages of the former method and none of its disadvantages, can be obtained by having a sufficient number of judges grade the questions for all the probable answers. With a table of such gradings for each question, it is then only a matter of a single grader comparing the answers of the pupils with the gradings of the tables. Anyone is then able to give and grade the test with an accuracy which is very nearly as great as the accuracy of the original grading of the expert judges. This second

method, which has been used, will be clearer when illustrated by reference to the specific questions.

Grading of the Interest Tests

Questions 1 and 2—*Magazines.*

Each of the magazines listed was graded from zero to ten for its significance as evidence of interest and information along the line of English and again with reference to history, by four judges (except that only three graded fifteen of the least familiar magazines), three of whom are psychologists and familiar with grading of this nature, and the fourth a librarian. The average, to the nearest integer, of the grades of the different judges is the grade given for each magazine in the following table:

Grade			Grade		
Eng.	*Hist.*		*Eng.*	*Hist.*	
1	0	1. All Story	3	1	37. Life
2	1	2. American Boy	5	2	38. Lippincott's
5	5	3. American	7	7	39. Literary Digest
1	0	4. Argosy	5	4	40. McClure's
10	4	5. Atlantic Monthly	1	0	41. Metropolitan
1	0	6. Black Cat	3	0	42. Modern Priscilla
0	0	7. Blue Book	1	0	43. Moving Picture World
8	2	8. Bookman			
2	1	9. Cassier's	1	1	44. Munsey's
7	4	10. Century	3	2	45. Musician
4	7	11. Collier's Weekly	5	9	46. National Geographic
6	9	12. Commoner	3	1	47. Outing
4	2	13. Cosmopolitan	7	8	48. Outlook
5	1	14. Country Life in America	3	2	49. Photographic Times
			2	0	50. Pictorial Review
5	1	15. Craftsman	2	1	51. Popular Mechanics
9	4	16. Current Literature	2	1	52. Popular Science Monthly
3	0	17. Delineator			
3	1	18. Electrical World	1	0	53. Printer's Ink
3	2	19. Etude	2	1	54. Puck
5	5	20. Everybody's	0	0	55. Red Book
4	2	21. Good Housekeeping	5	8	56. Review of Reviews
4	7	22. Graphic	6	1	57. St. Nicholas
0	0	23. Green Book	4	4	58. Saturday Ev'g Post
4	5	24. Hampton's	3	2	59. Science
4	6	25. Harper's Weekly	3	3	60. Scientific American
6	3	26. Harper's Monthly	7	3	61. Scribner's
4	5	27. Hearst's	3	0	62. Smart Set
3	0	28. Home Needlework	2	0	63. Strand
4	7	29. Ill'd London News	2	1	64. System
4	7	30. L'illustration	2	1	65. Technical World
4	7	31. Illustrirte Zeitung	1	0	66. Theatre
3	2	32. Industrial Eng'g	4	1	67. Woman's Home Companion
3	2	33. International Studio			
4	1	34. Ladies' Home Journal	3	4	68. Wilshire's
5	9	35. LaFollette's	5	8	69. World's Work
4	7	36. Leslie's Weekly	5	2	70. Youth's Companion

The correlations between the grades given by judges 1, 2, 3 and 4 are as follows:

r_{12}	.887
r_{13}	.893
r_{14}	.867
r_{23}	.788
r_{24}	.920
r_{34}	.821
Average	.863

Taking the number of judges for each magazine as three and one-half, the reliability coefficient equals .956.

The single grade given for these questions is the average grade for the magazines after altering the grade of the magazine marked "A" (usually about 1½ points) upon the basis of the answer to question 2, and weighting magazines A, B, C, D, and E 10, 8, 6, 4, and 2 respectively. The sample grading of the entire test, given on page 56, will show the steps in detail.

The reliability coefficient for the grading of each magazine is .956, but the reliability of the average of a number of such grades is higher. A factor tending in the other direction is the method of grading the magazine marked "A." The alteration of the grade of this magazine is not arbitrary, but left to the grader, and therefore its reliability is somewhat less than that of the magazines that have been graded by three or four judges. The net result of these two factors is probably to make the reliability of the grade given for the question about the same as the reliability of the grading of the magazines by the judges.

Question 3—*Books.*

The establishment of a guide for the grading of question 3 offers greater difficulties than was the case with question 1, for the reason that the number of books which may be preferred is unlimited. However, quite a number of books were repeatedly chosen, so that a grading for books for English and history covering 300 or so of the most frequent choices does cover a very large per cent of the books chosen. Furthermore, as each pupil chooses three books, it is much more than likely that two of the three will be books that are graded, or books by the same author as graded books, so that the grading actually is quite objective. The following is such a list, and the grades given, expressed as deviations from the mean, are the averages of the grades of from two to four judges, about one-half being graded by three or more judges. This list of about 300 titles is part of a larger list which included all the books preferred by the pupils. The larger list was given to four judges—two librarians and two others familiar with such work. The directions to the judges were to grade the books for English and history according to the following scheme:

Grade	English	History
1	The best literature.	Straight histories.
2	Excellent.	Books that are mainly historical. Historical biographies, etc.
3	Good.	Partly historical. Historical fiction.
4	Medium.	Fiction or adventure with traces of historical material.
5	Poor.	Adventure, etc., with no claim to any historical matter, but with lively action and plot.
6	Very poor—semi-trashy.	Non-historical fiction.
7	Pure trash.	Books with neither plot nor historical background, e. g., Electricity for Beginners.

The average correlation between the gradings of two judges is .882 for the English grading, and .720 for the history grading. The reliability coefficient, calling the number of judges $2\frac{1}{2}$, is, for the English grading .947, and for the history grading .861. Since all the books chosen are not in the following list, and since certain of those not so listed can be graded if the grader is familiar with the book or with the author, the reliability of the resulting grade of each book is somewhat less than the reliability coefficient just given.

The gradings from 1 to 7 of the different judges were combined into single grades, expressed as deviations from the mean,[1] and this is the grade given in the following table:

Author	Title	Grade	
		Eng.	*Hist.*
Scott	Abbot	1.2	1.2
Eliot	Adam Bede	2.0	−.2
Clemens	Adventures of Tom Sawyer	.2	.0
Doyle	Adventures of Sherlock Holmes	− .2	.0
Roosevelt	African Hunt	− .2	.0
Maeterlinck	Aglavine and Lysette	.9	− .8
Irving	Alhambra	1.4	2.2
Montgomery	Ann of Avonlea	− .4	− .8
"	Ann of Green Gables	− .4	− .8
Locke	Aristide Pujol	− .2	− .8
F. H. Smith	Armchair at the Inn	− .2	− .8
Franklin	Autobiography of Benjamin Franklin	.2	1.5
Stanley	Autobiography of Henry M. Stanley	.2	1.5
Haggard	Ayesha	−1.2	− .8
Wallace	Ben Hur	.5	1.2
M. Warde	Betty Books	− .8	− ,8
McCutcheon	Beverly of Graustark	−1.7	− .4
Thompson-Seton	Biography of a Grizzly	− .2	− .4
Wiggin	Birds' Christmas Carol	.1	− .8
Stevenson	Black Arrow	.8	− .6
Vance	Black Bag	−1.5	− .8
White	Blazed Trail	.2	.9
Dickens	Bleak House	1.2	− .2
Farnol	Broad Highway	− .6	− .4
London	Call of the Wild	.4	.5
Kipling	Captains Courageous	.4	− .4
Stockton	Casting Away of Mrs. Lecks and Mrs. Aleshine	.0	− .8
Garland	Cavanaugh, Forest Ranger	− .6	.4
Dumas	Chevalier de Maison Rouge	.5	1.2
"	Chicot the Jester	.5	.2
Dickens	Christmas Carol	1.3	— .8
Churchill	Coniston	.1	.6
L. Scott	Counsel for the Defense	− .3	− .8
Dumas	Count of Monte Cristo	.5	.8
"	Countess de Charny	.5	1.2
Gaskell	Cranford	.1	− .6
Barbour	Crimson Sweater	− .8	− .8
Churchill	Crisis	.1	1.2
"	Crossing	.1	1.2
Trowbridge	Cudjo's Cave	−1.7	− .4
Rostand	Cyrano de Bergerac	1.4	.9

[1] See Appx. p. 101.

Author	Title	Grade	
		Eng.	*Hist.*
Eliot	Daniel Deronda	1.4	.5
Stevenson	David Balfour	.8	.5
Dickens	David Copperfield	1.1	− .4
Ferber	Dawn O'Hara	− .6	− .7
Aguilar	Days of Bruce	− .3	1.1
Cervantes	Don Quixote	1.2	.9
Dickens	Dombey and Son	1.1	− .4
Major	Dorothy Vernon of Haddon Hall	− .6	.7
Bacheller	Dri and I	.1	.9
"	Eben Holden	− .2	− .5
Alcott	Eight Cousins	− .2	− .8
Tennyson	Enoch Arden	1.3	− .8
Longfellow	Evangeline	.4	.9
Drummond	Evolution of Man	.2	.4
Spenser	Faerie Queene	2.1	− .8
Poe	Fall of the House of Usher	1.2	− .8
Gaboriau	File Number 113	.1	.0
Fothergill	First Violin	− .4	− .8
Chaplin	Five Hundred Dollars	−1.5	− .9
Porter	Freckles	− .4	− .8
Read	Foul Play	.2	.0
Poe	Gold Bug	1.2	− .8
McCutcheon	Graustark	−1.7	− .5
Dickens	Great Expectations	1.1	− .5
Holmes	Gretchen	−2.7	− .8
Shakespere	Hamlet	2.3	.4
Dodge	Hans Brincker	− .8	− .5
Porter	Harvester	.1	.2
Thackeray	Henry Esmond	1.4	.9
Shakespere	Henry V	2.3	.9
Kelly	Her Little Young Ladyship	− .4	− .8
Redpath	History of America	.4	3.5
	Hollow of Her Hand	−1.4	− .9
Leblau	Hollow Needle	− .2	.0
Ford	Honorable Peter Sterling	− .4	− .2
Mulford	Hopalong Cassidy	− .9	− .8
Doyle	Hound of the Baskervilles	− .2	.0
Hawthorne	House of Seven Gables	1.2	− .8
Rohlfs	House of the Whispering Pines	−1.2	.0
Clemens	Huckleberry Finn	.2	− .5
Clemens	Innocents Abroad	.2	− .8
Davis	In the Fog	− .4	.1
Crawford	In the Palace of the King	− .2	.7
Deland	Iron Woman	− .4	− .8
Scott	Ivanhoe	1.2	1.2
McCutcheon	Jane Cable	−1.7	− .8
Bronte	Jane Eyre	.8	− .8
Ford	Janice Meredith	− .4	1.2
Craik	John Halifax, Gentleman	.2	− .8
Kipling	Jungle Book	.4	− .8

Author	Title	Grade	
		Eng.	*Hist.*
Scott	Kenilworth	1.2	1.2
Smith	Kennedy Square	− .4	− .8
Stevenson	Kidnapped	.8	.5
Kipling	Kim	.4	− .5
Irving	Knickerbocker History of New York	.8	.8
Williamson	Lady Betty Across the Water	−1.4	− .8
Scott	Lady of the Lake	1.2	.5
Cummins	Lamplighter	− .9	− .5
Lytton	Last Days of Pompeii	.8	1.2
Cooper	Last of the Mohicans	.4	.9
Reed	Lavender and Old Lace	−1.7	− .8
	Letters of Abraham Lincoln	.5	2.2
Johnston	Little Colonel in Arizona	− .9	− .8
"	Little Colonel at Boarding School	− .9	− .8
Kipling	Light that Failed	.2	− .8
Dickens	Little Dorrit	1.1	− .4
Alcott	Little Men	.1	− .8
Barrie	Little Minister	.6	− .8
Fox	Little Shepherd of Kingdom Come	− .4	− .2
Alcott	Little Women	.1	− .8
Blackmore	Lorna Doone	.2	.9
Parrish	Love under Fire	— .8	.2
Rice	Lovey Mary	− .6	− .8
Austen	Mansfield Park	.5	− .8
Rinehart	Man in Lower Ten	−1.0	− .5
London	Martin Eden	.5	− .8
Dickens	Martin Chuzzlewit	1.1	− .2
Bosher	Mary Cary	− .8	− .8
Zangwill	Master	.2	− .8
Corelli	Master Christian	−1.2	− .8
Eliot	Mill on the Floss	1.4	− .8
Marryat	Mr. Midshipman Easy	−1.2	− .8
Hugo	Miserables, Les	1.3	1.2
Lincoln	Mr. Pratt	−1.2	− .8
Barclay	Mistress of Shenstone	− .7	− .8
Rice	Mrs. Wiggs of the Cabbage Patch	− .6	− .8
Abbott	Molly Make-believe	− .6	− .8
Collins	Moonstone	− .2	− .8
Norris	Mother	− .4	− .8
	Mother Carey's Chickens	− .4	− .8
Spearman	Mountain Divide	− .8	− .6
Shakespere	Much Ado About Nothing	2.3	− .8
Verne	Mysterious Island	−1.2	− .4
Beach	Ne'er Do Well	− .9	− .4
"	Net	−1.2	− .8
Thackeray	Newcomes	1.4	− .6
Dickens	Nicholas Nickleby	1.1	− .2
Hugo	Notre Dame	1.3	.5
Homer	Odyssey	2.3	.9
Dickens	Old Curiosity Shop	1.1	− .4
Page	Old Gentleman of the Black Stock	− .2	− .8
Dickens	Oliver Twist	1.1	− .4
Burnham	Open Shutters	− .8	− .8

Author	Title	Grade	
		Eng.	*Hist.*
Darwin	Origin of Species	− .6	− .4
Alcott	Our Helen	.2	− .8
Dickens	Our Mutual Friend	1.1	− .4
E. Smith	Palace Beautiful	−1.2	− .8
Cooper	Pathfinder	.2	.7
Wells	Patty's College Days	−1.2	− .8
	Personal Memoirs of U. S. Grant	− .4	2.6
Dickens	Pickwick Papers	1.1	− .1
Cooper	Pioneers	− .2	.3
Flammarion	Popular Astronomy	− .9	−1.4
Austen	Pride and Prejudice	.4	− .8
Clemens	Prince and the Pauper	.2	− .8
Johnson	Prodigious Hickey	−1.2	− .8
Aldrich	Prudence Palfrey	.1	− .5
Harrison	Queed	.1	− .4
Scott	Quentin Durward	1.1	1.2
Sienkiewics	Quo Vadis	.4	1.2
Hornung	Raffles	−1.7	− .4
Jackson	Ramona	.4	.5
Wiggin	Rebecca of Sunnybrook Farm	− .2	− .8
Lytton	Rienzi	.8	1.2
Parkman	Robin Hood	− .8	.5
Thompson-Seton	Rolf in the Woods	− .2	− .4
Le Gallienne	Romance of Zion Chapel	.1	− .8
Eliot	Romola	1.9	.9
Barclay	Rosary	−1.2	− .8
Alcott	Rose in Bloom	.1	− .8
Hope	Rupert of Hentzau	−1.0	− .4
Hawthorne	Scarlet Letter	1.1	.3
Porter	Scottish Chiefs	− .4	−1.2
London	Sea Wolf	.5	− .4
Kipling	Seven Seas	.5	− .8
Haggard	She	−1.4	− .4
Wright	Shepherd of the Hills	.1	− .8
Goldsmith	She Stoops to Conquer	1.1	− .4
Eliot	Silas Marner	1.9	− .2
Porter	Song of the Cardinal	.1	− .8
London	Son of the Sun	.5	− .8
Barr	Souls of Passage	− .9	− .8
Reed	Spinner in the Sun	−2.2	− .8
Cooper	Spy	.4	1.2
Schreiner	Story of an African Farm	− .4	− .2
Keller	Story of My Life	− .2	− .8
Johnson	Stover at Yale	− .9	− .8
King	Street Called Straight	−1.0	− .8
Vaile	Sue Orcutt	− .8	− .6
Wyss	Swiss Family Robinson	− .7	− .2
Dumas	Taking the Bastile	.5	1.2
Dickens	Tale of Two Cities	1.1	1.2
Scott	Talisman	1.1	1.2
Jacobs	Texas Blue Bonnet	−1.0	− .8
Porter	Thaddeus of Warsaw	− .3	.8

Author	Title	Grade	
		Eng.	*Hist.*
	Thoughts of Marcus Aurelius	.7	− .1
Dumas	Three Guardsmen	.5	1.2
Jerome	Three Men in a Boat	− .4	− .8
Glyn	Three Weeks	−2.3	− .8
McCutcheon	Truxton King	−1.7	− .8
Verne	Twenty Thousand Leagues under the Sea	−1.2	− .8
Stowe	Uncle Tom's Cabin	− .2	.9
Ouida	Under Two Flags	− .4	.5
Henty	Under Drake's Flag	−2.4	.9
Alcott	Under the Lilacs	.1	− .8
Thackeray	Vanity Fair	1.4	− .2
Goldsmith	Vicar of Wakefield	1.1	− .8
Wister	Virginian	− .2	− .8
Scott	Waverley	1.1	1.2
Reed	Weaver of Dreams	−1.7	− .8
Malone	West Point Yearling	− .6	− .8
Doyle	White Company	− .2	.3
London	White Fang	.4	− .8
Thompson-Seton	Wild Animals I Have Known	− .2	− .4
King	Wild Olive	− .8	− .8
Scott	Woodstock	1.1	1.2
Rohlfs	Woman in the Alcove	−1.6	.0
Hawthorne	Wonder Book	1.1	− .8
Alger	Young Adventurer	−2.2	− .8
Bennett	Your United States	− .2	.3

The grade for question 3 is the average grade for the three books, weighting the first, second and third choices 3, 2, and 1 respectively.

Since this grade is an average of three grades, its reliability is greater than that of a single book, and is probably close to .94 for English and .86 for history.

Questions 4, 5 and 6—*Sports.*

The same general method used in establishing a guide for the grading of books, has here been used in drawing up a guide for the grading of sports, except that the grades were not expressed as deviations from a mean. The grading was on a basis of zero to ten, and the judges consisted of three persons experienced in such grading. The average correlation between the gradings of the different judges, when English and history are combined into a single correlation table, is .239. The reliability coefficient for the grading is, therefore, .485. The grading is separate for the preferences of boys and girls, as shown in the following table:

Sports	Boys			Girls		
	E	H	M	E	H	M
Motoring	3	3	2	3	3	2
Baseball	1	1	2	1	1	2
Basketball	1	1	2	1	1	2
Bicycling	1	1	2	1	1	2
Cards	1	1	8	....	....	....
Domestic activities (cooking, sewing, etc.)	....	....	....	3	2	1
Drawing, painting	4	4	3	3	4	3
Fishing	1	1	1	1	1	1
Football	1	1	1	....	....	....
Hockey	1	1	1	1	1	2
Horseback riding	2	2	1	2	2	1
Pool or billiards	0	0	2	0	0	2
Musical practice	2	2	2	2	2	2
Reading	9	9	2	9	9	2
Rowing or sailing	1	1	3	1	1	3
Shop work	2	1	9	....	....	....
Skating	1	1	3	1	1	3
Swimming	1	1	1	1	1	1
Tennis	1	1	1	1	1	1
Theatre	5	5	0	5	5	0
Running games	2	2	2	....	....	....
Walking	2	2	2	1	1	1
Watching a game	2	3	2	2	2	1

The grade for the question is the average, counting the indoor choice once, the outdoor choice once and the preference once. Since this grade is an average of three, one of them, however, a repetition, there is a slight tendency to increase the reliability of the grading over that of each sport singly. This tendency is probably counterbalanced by the fact that all possible sports are not included in the above list, and the grader must at times use his judgment in the matter. In view of these two facts, it is likely that the reliability of the grading is about .48.

Question 7—*Entertainments.*

The general method used in establishing a guide for the grading of magazines was followed here. Grading was upon the basis of zero to ten, and the judges consisted of four persons experienced in such work. The average correlation between the gradings of the judges, combining English and history into a single correlation table, is .929, so that the reliability coefficient of the grading is equal to .981. The grading follows:

Entertainments	E	H
Moving pictures	3	4
Circus	0	0
Football game	0	0
Baseball game	0	0
Track meet	0	0
Musical comedy	1	0
Vaudeville	0	0
Grand opera	7	4
Boxing contest	0	0
Band concert	2	1
Political rally	4	10
Light opera	3	0
Drama	9	7
Lecture, or stereopticon lecture on a subject that interests you	9	7

The grade given for the question is the grade of the entertainment marked "A," so that the reliability coefficient of this grade is .981.

Question 8—*Vocations.*

Here again, the number of choices is unlimited, so that the method used in the grading of sports is the one followed. The following list of vocations was drawn up after examination of all the choices of the pupils, and was graded by four judges. The average correlation between the gradings of the different judges, the mathematics, English and history gradings being combined into a single table, is .693, so that the reliability coefficient of the grading is .901.

Vocations	Boys			Girls		
	E	H	M	E	H	M
Actor	8	5	1	7	4	0
Artist	6	6	1	6	6	1
Athlete, professional	0	1	0	....	....	....
Author	10	7	2	9	7	2
Banker	6	5	8	....	....	....
Broker	5	4	6	....	....	....
Business	4	4	6	4	4	6
Carpenter	2	2	5	....	....	....
Decorator or designer	....	....	....	4	5	3
Doctor	6	5	3	6	5	4
Journalist	9	8	2	9	8	2
Engineering	4	4	9	....	....	....
Architectural	5	6	8	....	....	....
Electrical	4	3	9	....	....	....
Civil	4	4	9	....	....	....
Marine	4	4	9	....	....	....
Mechanical	4	4	10	....	....	....
Mining	4	4	9	....	....	....
Structural	4	4	9	....	....	....
Farming	1	2	2	1	2	2
Forestry	1	2	2	....	....	....
Housekeeper	....	....	....	2	2	1
Lawyer	8	8	2	7	6	2
Lecturer	....	....	....	9	8	1
Librarian	....	....	....	8	8	1
Milliner	....	....	....	1	2	1
Musician	3	3	2	3	3	2
Member of navy	4	7	5	....	....	....
Nurse	....	....	....	4	4	2
Politician	6	9	3	....	....	....
Secretary	....	....	....	8	7	5
Singer	4	4	2	4	4	2
Charity worker	5	6	4	6	6	4
Dressmaker	....	....	....	3	3	2
Teacher	8	8	8	8	8	7
Domestic science	....	....	....	6	4	4
English	....	....	....	10	8	4
Kindergarten	....	....	....	8	6	3
Physical training or dancing	....	....	....	5	4	2

Question 9—*Vocabulary.*

The grading of the words followed the general scheme used in grading the magazines. A number of words are graded zero by all the judges for their significance in one or all of the three lines, and as these do not enter into the grade which the pupil received, they have been omitted in calculating the reliability of the grading for mathematics, English and history, though simply as a measure of the reliability of the grading this is not desirable. The inclusion of these zero gradings would increase the reliability coefficient. The reliability differs considerably for English from that for mathematics and

history so that the coefficients are calculated separately. The average correlation for the different judges, excluding the zeros, as stated, between the gradings of the words, equals for mathematics .785; for English .542; for history .795, and the average for all is .707. Since three judges graded all the words, the reliability coefficients for the grades of these subjects are .916 for mathematics, .780 for English and .921 for history; an average of .872.

The grade the pupil receives for mathematics is the sum of the various words checked, after it has been multiplied by the factor of accuracy obtained in the last question; and similarly for English and history. Since this grade is made up of a considerable number of grades, whose reliability is given above, the reliability of the grade, before multiplication by the factor of accuracy, is appreciably greater than the reliability of the grading of each of the words, and is probably not less than .96 for mathematics and history, and not less than .92 for English.

As given in the next section, the reliability of the factor of accuracy is about .95, so that the reliability of the grade for the question is close to .91 for mathematics and history and in the neighborhood of .87 for English.

Question 10—*Factor of Accuracy.*

To determine the accuracy of the pupil's estimate of his knowledge, he is asked to define 13 of the words upon which he has previously expressed a judgment as to his familiarity with them. His definitions of these 13 words are graded 1, 2, 3, or 4. The sum of these grades gives a quantitative statement of the extent of the pupil's knowledge of the 13 words. This sum, which may be called the measure of the pupil's actual knowledge, divided by his claim, gives the factor of accuracy sought. In adding up the marks which constitute the pupil's own claim, it will be noticed that not infrequently the pupil has erased or marked over his previous marking, giving himself a lower mark in the second case. In all such cases the first grade put down by the pupil is the grade used. A magnifying glass may be of assistance, though it is seldom needed, the pupil making the correction probably not because he wishes to be dishonest, but because he realizes that he has over-estimated his knowledge and wishes to be honest and straighten it out.

The reliability of this factor of accuracy is undoubtedly high, and is estimated at .95.

Two Kinds of Reliability Coefficients

The reliability coefficients given, mean simply that the gradings of the questions as determined by the grader with the help of the tables formulated by the various judges, would probably correlate with the gradings determined by a second grader with the help of tables drawn up by a second set of judges, to the extent indicated. They do not mean that the gradings of the preferences of the pupils would correlate with similar gradings derived from different but similar data, to the extent indicated. The extent of this latter correlation can be determined for those questions that are not exhaustive, i. e., those for which a second similar test is possible. The vocabulary and the factor of accuracy questions fulfill these conditions. For the vocabulary test it would only be necessary to devise a second list of words, as similar as possible to the words in this list, have their significance evaluated by the same number of qualified judges as here used and have a second person grade them, using the guides of the judges. The correlation between the grades of the pupils in this second vocabulary test and their grades in the test here given, gives the reliability coefficient of the grading as a measure of the trait in question, based upon such a limited sampling as that here used.

Another method is to treat the halves of the present test as separate vocabulary tests and calculate the correlation between them. This latter method has been used for the history grading, giving a correlation of .620, derived from a sample of 36 pupils. The English and mathematics data are not so

extensive, so that the correlation in those cases would be somewhat smaller. Since the correlation between the gradings of the halves of the history words equals .620, the extent to which the two halves combined, or the whole test, would correlate with a similar test, is given by the usual formula $\frac{nr}{1+(n-1)r}$. This value equals .765, which is the reliability coefficient of the data as a measure of the pupil's vocabulary of historical words (this is, of course, before multiplication by the factor of accuracy).

Dividing the words used in determining the factor of accuracy, into two parts, one part consisting of words a, b, c, d, e, m and the other part of the balance, and proceeding in a similar manner, it is found that the reliability of the obtained factor of accuracy equals .453. The smallness of this reliability coefficient shows the limitation of the vocabulary test used, while the fact that the present test is significant, as will be shown later, demonstrates that a more accurate determination of the factor of accuracy would result in a vocabulary test of greater and very substantial value. Using the formula $\frac{nr}{1+(n-1)r}$ the reliability of a test, based upon any given number of words, may be obtained, and for certain numbers it is as follows:

Question similar to No. 9, except that number of words =	In which case the coefficient of reliability =
124	.867
186	.907
76	.80
Question similar to No. 10, except that number of words =	
26	.624
39	.713
52	.768
24	.60
37	.70
63	.80

It is apparent from these data that the determination of the factor of accuracy should be based upon about three times as many words as have been used, to make its reliability about the same as the grading of the history words.

Grade for Entire Interest Test

Mathematics:

The combination of the mathematics grades for the various questions into a single grade for the test, designated as M_i, is as follows:

$$M_i = 2(.2\,M_{Spts} + .05\,M_{Ents} + .4\,M_{Vocs} + 1.0\,F \text{ of } A + .08\,M_{Wds}).$$

The factor 2 is introduced to secure a better distribution.

M_{Spts} = grade of the sports for their mathematical significance, and similarly for other designations.[1]

English:

The single grade of the test for English is as follows:

$$E_i = .2\,E_{Spts} + .05\,E_{Ents} + .4\,E_{Vocs} - 5.0\,F \text{ of } A + .08\,E_{Wds} + 3.3\,E_{Mags} + .165\,E_{Bks}.$$

History:

The single grade of the test for history is as follows:

$$H_i = .2\,H_{Spts} + .05\,H_{Ents} + .4\,H_{Vocs} - 2.0\,F \text{ of } A + .08\,H_{Wds} + 3.3\,H_{Mags} + .165\,H_{Bks}.$$

Sample grading of the interest test:

The occasions for the exercise of judgment in this test are so few that a very few illustrations will suffice to make the method clear. To avoid decimals in the samples that follow, 5 times the measures contributing to M_i and 10 times those contributing to E_i and H_i are calculated.

[1] See Appx., pp. 101–103.

Name (Pupil No. 90). Date.....

Magazines Checked in Question I	Grade of Same for			Factor Contributing to Total Grade		
	M	E	H	M_i	E_i	H_i
× American		5	5			
× Atlantic Monthly		10	4			
× Blue Book		0	0			
× Bookman		8	2			
D Century (4)		28	16			
× Collier's Weekly		4	7			
× Cosmopolitan		4	2			
× Country Life in America		5	1			
C Craftsman (6)		30	6			
× Current Literature		9	4			
× Delineator		3	0			
× Electrical World		3	1			
× Etude		3	2			
× Everybody's		5	5			
× Good Housekeeping		4	2			
× Green Book		0	0			
× Hampton's		4	5			
× Harper's Weekly		4	6			
× Harper's Monthly		6	3			
× L'illustration		4	7			
× International Studio		3	2			
× Ladies' Home Journal		4	1			
× Leslie's Weekly		4	7			
× Life		3	1			
× Lippincott's		5	2			
× Literary Digest		7	7			
× McClure's		5	4			
× Metropolitan		1	0			
× Modern Priscilla		3	0			
× Munsey's		1	1			
A National Geographic (10) (See question 2)		50	90			
× Outing		3	1			
× Pictorial Review		2	0			
E Popular Mechanics (2)		4	2			
× Popular Science Monthly		2	1			
× Puck		2	1			
× Red Book		0	0			
× Review of Reviews		5	8			
B St. Nicholas (8)		48	8			
× Saturday Evening Post		4	4			
× Science		3	2			
× Scientific American		3	3			
× Scribner's		7	3			
× Smart Set		3	0			
× Strand		2	0			
× Technical World		2	1			
× Woman's Home Companion		4	1			
× World's Work		5	8			
× Youth's Companion		5	2			
× Outlook		7	8			
Frequency (75)		336	246			
Average grade		4.5	3.3			
33 times average grade					149	109

Answer to question 2: "The National Geographic Magazine is interesting because it tells so many interesting things of people we know very little, and of places none of us have seen." No warrant is found in this answer for altering the grade assigned to the National Geographic Magazine.						
Answer to question 3:						
1. Spenser's—Faerie Queene (3)		6.3	−2.4			
2. Scott's—The Abbot (2)		2.4	2.4			
3. Poe's—Fall of the House of Usher (1)		1.2	− .8			
Frequency (6)		9.9	− .8			
Average grade		1.67	−.13			
1.65 times average grade					3	0
Answer to question 4: "In a long walk through the woods."	1	1	1			
Answer to question 5: "Reading."	2	9	9			
Answer to question 6: "Number 4"	1	1	1			
Frequency (3)	4	11	11			
Average grade	1.3	3.7	3.7			
2 times average grade				3	7	7
Answer to question 7: "A" Grand Opera	0	7	4			
.5 times the grade				0	4	2
Answer to question 8:						
1st choice: "Landscape Architect (2)	16	10	12			
2d " "Designing."	3	4	5			
Frequency (3)	19	14	17			
Average grade	6.3	4.7	5.7			
4 times average grade				25	19	23

Answer to question 9:

Pupil's mark	Word						
0	simile						
3	primary election			15			
3	Mason and Dixon's line			27			
3	creed		6	6			
3	Acropolis		9	27			
1	rip saw						
3	hydrogen						
3	compound interest	18					
3	cube root	24					
0	paradox						
3	Saracens		6	30			
1	I. W. W.			6			
3	Whigs		6	27			
3	theosophy		9	3			
3	toga		6	27			
3	block plane						
0	NaCl						
3	fissure			3			

Pupil's mark	Word						
3	equation	21					
3	guillotine		9	27			
3	prose		15				
3	syndicalism			15			
3	H_2O						
3	transubstantiation		6	9			
3	gladiator		6	15			
3	debit	15		3			
3	gravity cell						
3	strata		3	3			
1	improper fraction	9					
3	lever						
3	ragtime						
2	physical valuation of R.R.			12			
3	score (in music)						
3	commercial fertilizer						
3	Magna Charta		6	30			
3	voucher	9		3			
0	ohm						
3	string halt						
3	fourth dimension	21					
3	piston rod						
0	Pythagorean proposition						
3	single tax			18			
3	stamen						
3	hemstitch						
3	Spanish Armada		6	30			
3	statute of limitations			15			
3	coherer						
3	vertebrate						
3	parallelogram	18					
3	omelette						
0	Reichstag						
3	Commerce Court			18			
3	states' rights			24			
2	space bar						
3	giblets						
0	Australian ballot						
2	mollusk						
3	perspective	3	3	3			
3	fireless cooker						
3	mortgagee	6		3			
3	referendum			15			
3	Formosa			12			
	Total	144	96	426			
	F. of A. (obtained in next question)						
	.57 × total	82	55	243			
	.8 times the grade				66	44	194

Answer to question 10:

Definitions of terms:	Grade given by pupil	Grade given by grader			
a. simile	0	0			
b. cube root "When a number is multiplied by itself three times."	3	2			
c. improper fraction	1	1			
d. rag-time—"music that is used in songs and vaudeville."	3	2			
e. physical valuation of railroads	2	0			
f. commercial fertilizer—"Fertilizer that is used in a commercial way."	3	0			
g. ohm	0	0			
h. Pythagorean proposition	0	0			
i. single tax	3	0			
j. hemstitch—"In sewing when threads are drawn out and then 4 or more drawn together in the centre."	3	3			
k. vertebrate—"The small parts that make up the back bone."	3	2			
l. parallelogram—"A figure where opposite sides are parallel."	3	3			
m. omelette—"A dish made from eggs."	3	2.5			
Sum	27	15.5			
Factor of accuracy $\frac{15.5}{27} = .57$					
$10 \times .57$			6		
$-50 \times .57$				-29	
$-20 \times .57$					-11
			100	197	324
M_i + mean (mean in question = 16)			20		
E_i + mean (" " " = 15)				20	
H_i + mean (" " " = 22)					32

or $M_i = 4$
$E_i = 5$
$H_i = 10$

In the following sample only the detailed grading is given where the judgment of the grader is involved:

Name (Pupil No. 14). Date.....

	Grade of Same for			Factor Contributing to Total Grade		
	M	E	H	M_i	E_i	H_i
Question 1.						
Combination of all the magazines, except that one marked "A."....(56)		160	83			
A, Life (See question 2)....(10)		30	30			
66		190	113			
Average grade.....		2.9	1.7			
33×average grade.....					96	56
Question 2.						
"I enjoy it because it is humorous and has some very interesting comments on the important things which are attracting attention."						
The interest here shown in current events warrants the increase in the history grading of Life, so, for this individual, the magazine is graded 3 for history instead of 1.						
Questions 3–9.						
Sum of the grades for questions 3 to 9 inclusive.....				59	39	135

Question 10: Terms and definitions of same	Grade for same given by: Pupil	Grader			
a. simile "A simile is a certain kind of a sentence."	1	1.5			
b. cube root—"I can't explain."	2	1			
c. improper fraction—"is a fraction which is not proper."	2	1.5			
d. ragtime—"A form of music with no special time."	3	1.5			
e. physical valuation of railroads—"Know nothing about it."	0	0			
f. commercial fertilizer—"Know nothing about it."	1	0			
g. ohm—"Can't explain."	1	1			
h. Pythagorean proposition	1	1			
i. single tax—"A tax on your personal belongings."	1	0			
j. hemstitch—"A certain kind of stitch. Can't explain."	1	1			
k. vertebrate—"Know nothing about it."	0	0			
l. parallelogram—"A figure. I can't explain."	2	1			
m. omelette—"Can't explain."	1	1			
	16	10.5			
Factor of accuracy = .66					
$10 \times .66$			7		
$-50 \times .66$				−33	
$-20 \times .66$					−15
			66	102	176
M_i + mean (mean in question = 18) =			13		
E_i + mean (mean in question = 17) =				10	
H_i + mean (mean in question = 26) =					18

or $M_i = -5$
$E_i = -7$
$H_i = -8$

The following sample covers determination of the factor of accuracy only:

Name (Pupil No. 158). Date.....

	Grade for Same Given by	
	Pupil	Grader
a. simile—"A comparison."	3	3
b. cube root—"A quantity multiplied by itself then into the product produces a certain cube."	2	3
c. improper fraction—"A fraction whose numerator is larger than its denominator."	1	3
d. ragtime	3	1
e. physical valuation of railroads—"Actual value of material and construction."	3	3
f. commercial fertilizer	0	0
g. ohm—"A certain degree to which magnets are wound."	3	2
h. Pythagorean proposition—"Proposition discovered by Pythagoras regarding squares over the sides of rectangular triangle."	3	3
i. single tax—"Tax on land only."	2	3
j. hemstitch—"An open kind of stitch used in sewing."	3	3
k. vertebrate—"An animal having a skeleton."	3	3
l. parallelogram—"A figure having two pairs of parallel sides."	3	3
m. omelette—"A preparation made with milk and eggs."	3	3
Factor of accuracy = 33/32 = 1.03	32	33

The grading of this interest test is not as long or difficult a task as it might at first seem. It can be greatly expedited by grading one question at a time after having memorized the table pertaining to that question, except the table for books, which it is impracticable to attempt to memorize. The use which is made of the interest test grades as well as the grades of the other tests, is given in the following section.

Combination of Grades of Various Tests for Purposes of Prognosis

Taking the grading of all the tests, there are six measures for each individual as follows: M_t, mathematics test, which is either the algebra or the geometry test; E_t, English test; H_t, history test; M_i, Mathematics interest test; E_i, English interest test; and H_i, history interest test. Not only do the gradings of each of these tests have significance in connection with the subject for which they are specifically graded, but they also have some significance for other courses. In other words, the most probable first-year grade in English may be said to be a function

of M_t, E_t, H_t, M_i, E_i, H_i. The regression equation, expressing E as a function of these six variables, might be calculated, but the labor would be very great, and therefore a slightly different method has been used, probably with little loss in the degree of correlation. The regression equation expressing E as a function of M_t, E_t, and H_t, is calculated, and this particular function is called E_{ct} (meaning the measure that represents that combination of the tests M_t, E_t and H_t, which correlates the most highly with English). A second regression equation expressing E as a function of M_i, E_i and H_i, is also calculated and designated as E_{ci} (meaning the measure which represents that combination of the interest tests M_i, E_i and H_i, which correlates most highly with English). Finally a regression equation is calculated expressing E as a function of E_{ct} and E_{ci}, and this function is designated as E_c (meaning the measure which represents that combination of E_{ct} and E_{ci} which correlates the most highly with English). So far as English is concerned, the entire object of the tests has been the derivation of this measure, E_c; and the correlation between E and E_c establishes the extent to which grades in the tests given serve as a basis for the prognosis of ability in high school English. The same procedure is followed with reference to mathematics and history, leading to measures M_c and H_c. The following sections will be devoted to explaining the derivation of M_{ct}, E_{ct}, H_{ct}, M_{ci}, E_{ci}, H_{ci}, M_c, E_c and H_c, in the order named.

M_{ct}—*Combination of Tests with Reference to (a) Algebra and (b) Geometry*

(a) Algebra: In order that it may not be lost sight of, it is repeated here that all of the measures mentioned in the last section are measures that are expressed as deviations from the means of the groups to which the measures belong. For one duplicating this test, the means given on page 68 may be assumed, or better, they may be calculated anew for the group tested.

The combination A_{ct}, of A_t, E_t, H_t, which correlates the highest with A is as follows:[1]

$$A_{ct} = .6A_t + .4E_t + .11H_t$$

This equation is self explanatory. To obtain A_{ct} it is only

[1] See Appx., p. 99.

necessary to add .6 of A_t, .4 of E_t and .11 of H_t, paying proper attention to sign. The correlation between A and A_{ct} equals .48. The apparent weighting of the three tests, .6, .4, .11 is not the exact weighting of the tests, for the standard deviations affect these regression coefficients. Since the standard deviation of H_t is large the weighting is somewhat greater than .11. The weighting seems very reasonable, bearing in mind that H_t is not as reliable a measure as E_t as it has but a single measure entering into it, whereas E_t is an average of three.

(b) Geometry: The regression equation for geometry is:

$$G_{ct} = .8\,G_t + .08\,E_t + .184\,H_t$$

The correlation between G and G_{ct} equals .43. The small weighting of E_t is somewhat of a surprise. It must be assumed that some of the elements entering into the grading E_t are more directly related to algebra than to geometry.

The measures A_{ct} and G_{ct} are entered in the same correlation tables and designated as M_{ct}, for purposes of determining the relative weighting of mathematics, English and history tests.[1]

E_{ct}—*Combination of Tests with Reference to English*

That combination of M_t, E_t, H_t which proved the most feasible is:[2]

$$E_{ct} = \tfrac{1}{3}(M_t + E_t + H_t).$$

The correlation between E and E_{ct} equals .46. The weighting here used yields a correlation practically as high as that given by the regression equation but this is not an accurately determined regression equation and it is impossible to use it for determining the relative importance of the factors M_t, E_t and H_t with reference to E. It may, however, be said that they do not differ greatly in their relative bearing upon E.

H_{ct}—*Combination of Tests with Reference to History*

The combination for history is the same as that for English:[3]

$$H_{ct} = \tfrac{1}{3}(M_t + E_t + H_t)$$

[1] See Appx., pp. 99, 105.
[2] See Appx., p. 100.
[3] See Appx., p. 100.

No accurate analysis of the importance of the factors M_t, E_t and H_t with reference to their bearing upon history is possible from this datum.

M_{ci}—*Combination of the Interest Tests with Reference to Mathematics*

The regression equation giving that combination of M_i, E_i and H_i which correlates the highest with M is:

$$M_{ci} = .5\,M_i + .65\,E_i - .2\,H_i$$

The correlation between M and M_{ci} equals .30. The mathematics interest test is weighted the most heavily in this equation in spite of the fact that the coefficient of E_i is the largest. This comes about from the fact that the standard deviation of E_i is considerably smaller than that of M_i. The actual weighting of the different elements is approximately in the ratio of 224: 183: −95. The occasion of the negative weighting of the history interest test may be determined from the raw data for the calculation of the regression equation given in the Appendix.[1] In brief it is due to the low correlation between M and H_i, .15, and the high correlations between M_i and H_i, .54, and between E_i and H_i, .63. Why the first of these three correlations is low and the second and third high is not apparent—an accurate calculation of the regression equation involving the parts of the interest test would reveal the cause, but it would be a very laborious task.

E_{ci}—*Combination of the Interest Tests with Reference to English*

The regression equation giving that combination of M_i, E_i, and H_i which correlates the highest with E was found to have only a trifling advantage over the use of E_i alone.[2]
Therefore the relation used is:

$$E_{ci} = E_i$$

The correlation between E and E_{ci} equals .46.

---

[1] See Appx., pp. 103–104.
[2] See Appx., p. 104.

H_{ci}—*Combination of Interest Tests with Reference to History*

The regression equation giving that combination of M_i, E_i, and H_i which correlates the highest with history is:[1]

$$H_{ci} = -.5\,M_i + .38\,E_i + .7\,H_i$$

The correlation between H and H_{ci} equals .33. The negative weighting of H_i in its bearing upon mathematics is comparable to the negative weighting here, of M_i in its bearing upon history. The data for definitely determining the cause of this latter, as well as the former, are lacking.

M_c—*Combination of* M_{ci} *and* M_{ct} *with Reference to Mathematics*

The regression equation giving that combination of M_{ci} and M_{ct} which correlates the highest with mathematics is:[2]

$$M_c = .66\,M_{ci} + 1.00\,M_{ct}$$

The correlation between M and M_c differs somewhat for the pupils taking geometry from that for the algebra pupils. The correlations are: $r_{GG_c} = .44$, $r_{AA_c} = .49$. These correlations are not as high as could be desired, nor as high as the correlation between grammar grade mathematics and first year mathematics, which is .58. However, r_{AA_c} is only .09 less than $r_{F_M(7,\,6,\,5,\,4_M)}$, and when it is considered that the former is a correlation based upon tests of a few hours duration while the latter is based upon the work of four years, it is a very satisfactory showing and is of positive value for purposes of prognosis and classification.

Lacking information as to the pupil's past performance, classification, at present, usually depends upon such things as the pupil's, or teacher's, preference as to the hour when the subject is to be taken, or upon the first letter of his last name, or some other equally irrelevant point. It is earnestly hoped that tests will be devised enabling a very accurate prognosis, but, pending such tests, there is nothing to lose and everything to gain by the use of the tests here given, whose significance has been accurately evaluated upon the basis of the performance of some 235 pupils. A summarized statement of the procedure in using the test data is given on page 68.

---

[1] See Appx., pp. 104–105.
[2] See Appx., p. 105.

E_c—*Combination of* E_{ci} *and* E_{ct} *with Reference to English*

E_{ci} and E_{ct} have equal significance in determining the most probable standing in English. The regression equation is:

$$E_c = E_{ci} + E_{ct}$$

The correlation between E and E_c is equal to .55. This correlation is higher than the correlation just above, for mathematics, though considerabley less than $r_{F_E(7, 6, 5, 4_E)}$, which equals .71. It is interesting to note that whether dealing with grammar school grades or with special tests it is possible to give a closer estimate of a pupil's performance in English than it is in mathematics. The greater difference between the natures of algebra and arithmetic, than between high school English and elementary school English, is probably a contributing cause in the case of these special tests as well as in the case of elementary school records.

H_c—*Combination of* H_{ci} *and* H_{ct} *with Reference to History*

The regression equation giving the best combination of H_{ci} and H_{ct} with reference to history is:

$$H_c = .4\,H_{ci} + H_{ct}$$

The correlation between H and H_c equals .49. The apparent unimportance of H_{ci} in comparison with H_{ct} exists only in part, as the standard deviation of H_{ci} is much larger than that of H_{ct}. The actual relative weighting of H_{ci} and H_{ct} is approximately in the ratio of 22:39.

Use of Regression Equations

Substitution of the test grades in the regression equations is required in order to use them for purposes of estimation of probable high school standing in the subjects, algebra, geometry, English and history. After having given and graded the tests for the pupil whom it is desired to examine the grades are substituted in the equation giving A_c, for purposes of estimation of his probable standing in freshman algebra, in the equation giving E_c, for estimation of his probable English standing, etc. The necessary equations are here summarized:

$M_c = A_c$ or G_c.

$A_c = .66\,M_{ci} + M_{ct} = .33\,M_i + .44\,E_i - .132\,H_i + .6\,A_t + .4\,E_t + .11\,H_t$.

$G_c = .66\,M_{ci} + M_{ct} = .33\,M_i + .44\,E_i - .132\,H_i + .8\,G_t + .08\,E_t + .184\,H_t$.

$E_c = E_{ci} + E_{ct} = E_i + \frac{1}{3}(M_t + E_t + H_t)$.

$H_c = .4\,H_{ci} + H_{ct} = -.2\,M_i + .15\,E_i + .28\,H_i + \frac{1}{3}(M_t + E_t + H_t)$.

$M_{ci} = A_{ci}$ or $G_{ci} = .5\,M_i + .65\,E_i - .2\,H_i$.

$E_{ci} = E_i$.

$H_{ci} = -.5\,M_i + .38\,E_i + .7\,H_i$

$M_{ct} = A_{ct}$ or G_{ct}.

$A_{ct} = .6\,A_t + .4\,E_t + .11\,H_t$.

$G_{ct} = .8\,G_t + .08\,E_t + .184\,H_t$.

$E_{ct} = \frac{1}{3}(M_t + E_t + H_t)$.

$H_{ct} = E_{ct}$.

$M_i = A_i$ or $G_i = \frac{1}{5}(2\,M_{Spts} + .5\,M_{Ents} + 4\,M_{Vocs} + 10\,F$ of $A + .8\,M_{Wds} -$ mean).

$E_i = .1\,(2\,E_{Spts} + .5\,E_{Ents} + 4\,E_{Vocs} - 50\,F$ of $A + .8\,E_{Wds} + 33\,E_{Mags} + 1.65\,E_{Bks} -$ mean).

$H_i = .1\,(2\,H_{Spts} + .5\,H_{Ents} + 4\,H_{Vocs} - 20\,F$ of $A + 8\,E_{Wds} + 33\,E_{Mags} + 1.65\,E_{Bks} -$ mean).

$M_t = A_t$ or G_t.

$A_t = \frac{1}{5}$(Sum of grades of all the problems − mean).

$G_t = \frac{1}{3}$(Sum of grades of problems 1, 7, 8, 9, 10 − mean).

$E_t = \frac{2}{3}(E_a + E_v + W -$ mean).

$H_t = 2(H_a -$ mean).

To obtain any of the last eight grades subtract the mean from the pupil's gross mark and divide the remainder by the divisor given in the equations and repeated in the following table of means and divisors.

Group and Class	Math. Test M_t		Eng. Test E_t		Hist. Test H_t		Math. Int. M_i		Eng. Int. E_i		Hist. Int. H_i	
	Means	Dvsrs.	Means	Dvsrs.	Means	Dvsrs.	Means	Dvsrs.	Means	Dvsrs.	Means	Dvsrs.
1 Beginning 2d year	G_t25	3	18.5	3/2	7	1/2	95	5	170	10	260	10
2 Mid-year 1st year	A_t89	5	16	3/2	6	1/2	80	5	150	10	220	10
3 Entering 1st year	A_t72	5	15	3/2	6	1/2	65	5	160	10	200	10
4 Beginning 2d year	G_t26	3	19	3/2	5.5	1/2	100	5	190	10	280	10
5 Entering 1st year	A_t88	5	17.5	3/2	6.5	1/2	70	5	170	10	240	10

If $\bar{A}$ represents the most probable algebra grade expressed as a deviation from the mean, for some given numerical system of marking it is related to A_c as follows: $\bar{A} = r_{AA_c}\frac{\sigma_A}{\sigma_{A_c}}A_c = .49\frac{\sigma_A}{3.736}A_c$ in which A_c is given for each pupil by the tests and σ_A is to be determined for the school in question. If the grading is on a percentile scale with a passing mark of 70, σ_A will be in the neighborhood of 8. The following are the standard deviations of the type σ_{A_c}.

$$\sigma_{A_c}=3.736;\ \sigma_{G_c}=4.066;\ \sigma_{E_c}=5.328;\ \sigma_{H_c}=3.622.$$

If $\bar{G}$, $\bar{E}$, and $\bar{H}$ have similar meanings to $\bar{A}$, their values are as follows:

$$\bar{G}=.44\frac{\sigma_G}{4.066}G_c;\ \bar{E}=.55\frac{\sigma_E}{5.328}E_c;\ \bar{H}=.49\frac{\sigma_H}{3.622}H_c.$$

The calculation may be presented as follows:

Gross Grade of Pupil Group 3		Means	Deviations from Mean	Divisors	Deviations ÷ Divisors
Alg. test	71	72	−1	5	$0=A_t$ or M_t
Eng. test	22.5	15	7.5	3/2	$5=E_t$
Hist. test	9	6	3	1/2	$6=H_t$
Math. int. test	68	65	3	5	$1=A_i$ or M_i
Eng. int. test	153	160	−7	10	$-1=E_i$
Hist. int. test	141	200	−59	10	$-6=H_i$

Substitution of these values in equations M_c, E_c, H_c gives $A_c=4$, $E_c=3$, $H_c=2$. If these measures are divided by their respective standard deviations ($\sigma_{A_c}=3.736$, $\sigma_{E_c}=5.328$, $\sigma_{H_c}=3.622$), the resulting measures, 1.070, .563, .552, express the predicted standing in terms of the standard deviation.

The calculation carried through is typical, being in fact that of pupil No. 157, and it is to be noted that the prediction is quite selective, for the estimate for algebra, 1.070, is one-half a standard deviation larger than the prediction for English. The advantage of such a differential prognosis over one of average ability only is very evident, if classification of the pupil is the aim. The extent to which the method performs this differential diagnosis may be measured by taking the differences between measures such as 1.070 and .563 and correlating them with the actual differences in grades in mathematics and English courses,

these latter likewise expressed in units of the standard deviations. This correlation may be expressed by the symbol $r_{(M-E)(M_c-E_c)}$ and calculation gives its value to be .31. This correlation is far from negligible and indicates the selective nature of ability. It may be expected that further use of the same method will lead to higher correlation and to very valuable tests for purposes of differential diagnosis.

The scheme just given for evaluating the record of any one pupil is a rather long process. It will be found that there is a very great saving in time if the work is tabulated and the steps performed one at a time upon the entire number of pupils' records, rather than as in the example, where the steps were performed consecutively and in toto for one pupil before going on to the next pupil.

The time of giving and evaluating the tests can be further shortened by omitting those parts not used in the final result. In the geometry test the following problems may be omitted: 2, 3, 4, 5, 6. In the English and history tests omit the grading for dramatization, and in the history test omit the grading for valuation. It is, however, recommended that in the interest test a larger number of words be used in determining the factor of accuracy.

It will at times happen that test records for a given pupil are incomplete, in which case a zero may be put in for the missing record, without greatly lessening the significance of the measure. It would be a little more reliable to estimate closely, from the data at hand, the probable value of the missing record and enter it, but there is danger that the estimate will be quite inaccurate, in which case it is worse than no estimate at all. The assigning of a zero grade is simply assigning the mean grade.

SECTION 6.—USE OF ALL SOURCES OF DATA IN ESTIMATING PROBABLE AVERAGE STANDING

Of the group of pupils whose elementary school records were available 33 were entering first year of high school. All of the following data were available with reference to these 33: (1) average first-year grade, F_A; (2) elementary school grade (7, 6, 5, 4_A); (3) teachers' estimates, Est_A; (4) special tests M_c, E_c, and H_c. The special tests are arbitrarily combined into a single measure by averaging,

$$\frac{M_c+E_c+H_c}{3}=T_A.$$

To determine the total bearing of these four sources of data upon average first-year standing the regression equation combining them with reference to it has been calculated from the data in the accompanying table.

	F_A	$(7, 6, 5, 4_A)$	Est_A
$(7, 6, 5, 4_A)$	.83		
Est_A	.81	.68	
T_A	.51	.56	.54

It is to be noticed that here $r_{F_A(7, 6, 5, 4_A)}=.83>.789$ (given on p. 8), and that $r_{F_A Est_A}=.81>.76=r_{Av(I.a.,'Cons.,Emo.i.,Exp.)}$. (Given on p. 16). These differences are probably entirely accounted for by fluctuations due to sampling. The smaller values are the more reliable, being based upon larger populations.

The regression equation is [1]

$$F_A=.536\ (7,\ 6,\ 5,\ 4_A)+.481\ Est_A-.043\,T_A.$$

The negative regression coefficient,—.043, is probably due to fluctuations in sampling. The probable error of the partial correlation coefficient entering into this regression coefficient is .112, so that no great significance can be attached to its negative value.

$r_{F_A[(7, 6, 5, 4_A),Est_A T_A]}=.89$, with a probable error of .023. This very high correlation is of interest in showing the great stability of individual character. To know that a pupil's grades

[1] See Appx., p. 106.

in the first year of the high school are so largely determined by what he possesses within his own personality is convincing evidence of the paramount importance of nature over and above nurture. With the undoubtedly varying environments under which these pupils lived there would be a greater divergence between estimate and accomplishment if nurture were the major factor.

There is a likelihood that the correlation of .89 is higher than might be ordinarily expected from similar data, for, as noted, the particular sample dealt with seems to show a slightly closer relation than usual. This is true to the greatest extent in the case of teachers' estimates, where it is likely that the teachers who made the estimates were particularly well acquainted with the pupils, for these pupils had been in the elementary department of the same school for the preceding four years, and their capacities for accomplishment were probably very well known. The lack of absolute independence between teachers' estimates and average class standing, noted on page 15, is a factor to be borne in mind.

In case it is essential to obtain as close an estimate as possible of a pupil's ability all three sources of data could profitably be used, but for ordinary needs of classification one source should be adequate. The method of combining the three measures into a single measure is given in the Appendix, p. 106.

SECTION 7.—THE AGE OF PUPILS AS A FACTOR

The number of factors involved in this study has been so great that the age factor has been omitted.

The correlation between average class standing and age, using all the data, is −.31. Eliminating the bearing of innate mental capacity (or mental capacity as existing at one certain age) would certainly give a positive partial correlation between age and standing in a given grade. Though it is a fact that the average twelve-year-old first-year pupil has a higher average standing than the average sixteen-year-old first-year pupil, it of course is not true that the average twelve-year-old is brighter than the average sixteen-year-old.

The occasion of the negative total correlation, −.31, is probably due to the fact that dull and over-age pupils are advanced more rapidly than their talents warrant, thereby always keeping them in a class which taxes their capacities and in which they can secure only low marks.

Since there exists this negative correlation between age and average standing in a given grade, the use as a measure of intelligence, of the age at which a pupil reaches a certain grade, gives to the bright pupil but part of the credit due him. The bright pupil is less advanced, and the dull pupil more advanced, judged by the grade attended, than talent warrants. The effect of this is to make the measure "age of attaining a certain grade" less reliable as a measure of ability than it otherwise would be.

SECTION 8.—COMPARISON WITH OTHER STUDIES

The most fundamental distinction between this and the great majority of correlation studies is that the aim of this study is prognosis and not at all to establish the existence and magnitude of some theoretical relationship. This fact has already been referred to but is mentioned again as some important points of method depend upon it. It will be noted that, in the following paragraphs, where comparison with other investigations of mental relationships is impossible, it is generally due to difference in method, necessitated by this difference in purpose.

A number of investigations, notably several by Spearman and his pupils, have as their object the determination of the abstract relationship which exists between certain tests and mental capacities. The aim is to establish the relationship that exists after errors of sampling, observation, and the like have been eliminated. The direct conclusions from such studies are of necessity theoretical, whatever may be the indirect practical implications resulting therefrom. Acting upon these implications educational practice might be altered, but it still would remain to be seen if it were bettered thereby.

This is not a criticism of theoretical investigations for, by suggesting relationships and methods, they have been the forerunners of progress; but it is for the purpose of pointing out the difference in object and the consequent difference in method between such investigations and the present one, which has as its object the utilizing of measures obtainable under ordinary class-room conditions, with whatever errors may be inevitable, for whatever they actually demonstrate themselves to be worth as evidence of the capacity it is desired to measure.

If one set of test measures correlates with class standing to a certain extent, no amount of superimposed treatment for elimination of observational errors, chance errors, or the like can change this raw relation which exists. Correction for attenuation would lessen the accuracy and vitiate the significance of the use of marks as a prognosis of other "raw" marks. Knowing the correlation between tests and average standing and wishing

to estimate the latter from the former, the most reasonable prediction is that given by the regression equation. This is exactly what the regression equation has been devised to give, and "correction" of the correlation coefficients in any way at all would lead to less accurate estimation. For this reason, none of the studies, the conclusions of which are based upon "corrected" coefficients of correlation, are comparable with this work, nor should the size of the coefficients of correlation here obtained be compared to "corrected" coefficients. The latter are meant to be a prophecy of what would be the correlation provided errors of various kinds were absent, while the former state the relation between existent measures.

The results of this study, however, do shed some light, and give a method of attack, upon the problem of the existence, or non-existence, of a single mental function which is paramount in all intellectual activities. The statement of the view of those holding to the idea of a single mental function has undergone much development and elaboration, until it now seems to be about as follows:—that every intellectual performance depends not only upon a general factor, "but also in varying degrees upon a factor specific to itself and of very similar performances."[1]

How the most ardent advocate of the specific nature of ability can object to such a statement, is hard to see. The problem is no longer a qualitative but a quantitative one. It is now necessary to measure intellectual performances and ascertain what part of each is a common element and what part is unique. The regression equation method, involving more than two variables, is beautifully adapted to solving this problem, and correlation between differences in accomplishment, or capacity, gives first-hand testimony as to the uniqueness or generality of mental function.

Pupils in the elementary school demonstrate a unique ability along the line of mathematics or English, by getting, relative to their average accomplishment, higher or lower marks in these subjects. That these marks do represent a unique ability and are not due to chance and the vicissitudes of teachers' gradings, is evidenced by the fact that different teachers, in the high

[1] See B. Hart and C. Spearman, General Ability, Its Existence and Nature, *British Journal of Psychology*, 1912. Also C. Spearman, Theory of Two Factors, *Psych. Rev.*, Vol. 21, No. 2.

school, recognize the same relative superiority or inferiority in the one subject, or other. The correlation $r_{(F_{M-E})(7, 6, 5, 4_{M-E})} = .52$, with a probable error of .065. The size of the probable error precludes the possibility of the correlation being due to chance. The alternative is that intellectual function is specific, unless it is argued that ability to secure grades is not solely an intellectual function.

This might be maintained, but grades have been used by proponents of the general factor theory as measures of intellect; and, furthermore, if so fundamental a mental characteristic as the ability to earn grades is not a fit capacity for consideration in connection with the general factor theory, then the theory must be of very limited scope in its application, and the traits of importance for scholastic and business success will lie outside its realm. The same conclusion may be drawn just as convincingly from the correlation $r_{(M-E)(Mc-Ec)} = .31$, for its probable error is only .040.

Another requirement of method is that the means used in the study shall be capable of determination at the time of prognosis. Studies which have dealt with the correlation between high school and university marks, or between elementary school and high school marks, have, without exception so far as the author is aware, selected the group upon the basis of attendance in the higher school, and then calculated the means in the lower school of the group thus selected. It follows that the means used for the lower school data are not capable of determination until the selection has occurred upon the basis of attendance in the higher school. Any elimination that takes place is entirely obscured by the method, and the use of the correlation found, for purposes of prediction, is not sound because it is not known from what mean, deviations should be measured.

However, though theoretically justified this criticism is probably not of very great moment when dealing with elementary school and first-year high school pupils. The evidence of this study is that there is not a sufficient selection of the brighter pupils in passing from the elementary school to the high school to necessitate changing the elementary school mean of pupils who attended high school from the mean of elementary school pupils in general. It is quite possible that this would not be

true in schools where there is a greater elimination than in the well-to-do schools from which these data are obtained.

The study of Dearborn[1] is excellent evidence that high school efficiency is highly correlated with university proficiency, but the method is not a serviceable one for a quantitative prognosis problem; and the high school means in his distributions are means of high school pupils who later attended college and are therefore the means of a selected group.

The same remark may be made in regard to the means used in the study by Miles[2] and it may be a material point in this case, for the amount of elimination between the elementary school and the fourth year of the high school is very much more extensive, and probably also selective, than between the elementary school and the first year of the high school. Miles finds that the correlation between the average elementary school grade and the average high school grade is .71. This is quite in harmony with the results of the present study and it is probable that Miles' data, treated by the regression equation method, would yield correlations between .80 and .90.

The fact that Miles deals with the average of all high school grades results in higher, or lower, correlations than would be obtained in dealing with first-year high school marks only, dependent upon which of the two following factors is the stronger: (1) In general, as the time between testing is increased the correlation decreases; and as the second, third and fourth years of the high school are more and more remote from the elementary school in time, it might be expected that correlation between the elementary school record and the first-year record would be greater than that between the elementary school record and the average of the entire high school record. (2) A factor tending to offset this is the fact that the reliability of an average increases as the number of grades entering into it increases, and, to the extent that a grade represents native ability, the greater the number of grades averaged the greater the reliability of this measure. It is impossible to say, *a priori*, which of these factors is the more important, but it is the author's opinion that the

---

[1]See Dearborn, W. F., Relative Standing of Pupils in the High School and in the University, *Wis. Univ. Bulletin*, No. 312, 1909.

[2]See Miles, W. R., Comparison of Elementary and High School Grades, *Univ. of Ia. Studies in Education*, Vol. 1. No. 1.

importance of the second factor has been, quite generally, undervalued, and might easily be the more important of the two.

Another class of studies has been undertaken, especially in England by investigators who have found the correlations between various tests and intellectual ability—the latter based upon teachers' and headmasters' estimates. It is pertinent to ask what relation there is between intellectual ability and the ability to secure grades.

The regression equation giving the bearing of teachers' estimates of intellectual ability, conscientiousness, emotional interest, and oral expression, upon average class standing, weights these factors in the ratio of 8:4:2:1, or, combining the first and last and designating it as the intellectual factor, and combining the second and third and designating it as the motive factor, or factor of effort, it is seen that the weighting is in the ratio of 3:2. Since effort is so important a factor in accounting for the ability to secure grades, it is apparent that the correlation between tests and intellectual ability will be quite different from, and probably higher than, the correlation between the same tests and class standing. This is a common finding and in the study by Wyatt[1] it is possible to estimate the extent of this difference.

Wyatt finds that the average correlation between his tests and intelligence, as determined by the headmaster's estimate, averaged .63, and that the correlation, for a different group, between the same tests and intelligence, as judged by class standing, averaged .51. As the headmaster did not grade upon both intellectual ability and effort it is probable certain evidence of excellent effort received credit as intellectual ability. Accordingly it may be expected that the difference in correlation between Wyatt's tests and real intellectual ability, from that between his tests and class standing, is actually greater than the .12 found.

Wyatt's tests apparently were given at about the same time that the marks which determined class standing were earned, so that his results are not comparable with the results of this study. Also the age of pupils is different, but his results suggest that certain of the tests used, especially the analogy and completion tests, are highly indicative of average class standing, and tests

[1]See Wyatt, S., Quantitative Investigation of Higher Mental Processes, *Brit. Jour. of Psyc.*, Vol. VI, Pt. 1.

of this nature are worthy of investigation for purposes of estimating average capacity, but it is doubtful if they have particular value for purposes of differential prognosis. In giving such tests it is not to be expected that the care with which Wyatt gave them will be duplicated under ordinary class conditions.

There are at least two classes of tests which are comparable with the tests here given, so far as purpose is concerned. One of these is entrance examinations. They perform their task, in the main, by attempting to measure acquired knowledge, whereas the tests here given, in the main, attempt to measure interest and capacity. Both types of examinations have a function to perform and the former should be supplemental to the latter in the final determination of the classification of the pupil. Acquired knowledge tests, of themselves and alone, are too likely to be evidence of the degree of success which has attended a cramming process, and not very definitely evidence of ability, which is the more important consideration. The following correlations, given by Thorndike,[1] show a progressive decrease in correlation between the median entrance examination grade and the average grade in the different years of the college course; freshman year .62, sophomore year .50, junior year .47, senior year .25. Intellectual capacity could hardly have changed much, relatively from pupil to pupil, during the four years of the college course. These correlations seem to indicate that the capacity measured by the entrance examination was, in the main, acquired knowledge and not intellectual ability, otherwise, why the decrease from year to year?

For purposes of immediate differential diagnosis tests of acquired knowledge undoubtedly perform an important function, but for the broader problems of vocational guidance and the selection of general courses of study they have very limited scope.

The second class, tests of the Binet type, have classification as their object, and in this respect are comparable to the tests here given. Thus far, however, they have not shed much light upon the points of relative strength or weakness of the individual tested. If mental deficiency is not general, but selective, an individual being normal in one capacity and quite defective in

[1] See Thorndike, E. L., in *Science*, N. S., Vol. XXIII, p. 839.

a second, then a mass test of mental age gives no light upon the distinctive feature which it is desirable to be acquainted with. There is plenty of evidence to indicate that deficiency is selective in many cases and the defect of the Binet tests on this point should be remedied. However, there is sufficient correlation between defects to make a Binet test of considerable value for purposes of classification; just what value has never been determined, so far as the author is aware, in quantitative terms, i.e., in terms of the correlation between capacity, as estimated from the test, and capacity as determined by as complete and conclusive measurements as possible. It is essential that mental age tests be tested by such methods, in order to judge which are the more accurate and what their accuracy is.

The use of tests of this nature as a guide to classification may be illustrated by the work of Adler, in New York School 77.[1] Boys in the first and fourth grades were tested with Dr. Goddard's 1911 revision of the Binet tests, with additions from the tests of Terman, Whipple and Courtis. In both the first and fourth grades the 35 pupils, out of about twice that number, who tested highest, were placed in an advanced class. The results were highly satisfactory. To quote the results in the case of the advanced section of the fourth grade: "Twenty-two of the thirty-five pupils are ready to begin the second half of the fifth grade work. Thirteen of the pupils begin the regular fifth grade work, though several of these will probably catch up with the advanced pupils before the end of the term. One pupil, who was absent because of contagion, will be retarded." The tests are evidently of high significance, but the calculation of the coefficient of correlation between them and the accomplishment in class would be of value in giving a quantitative measure to the degree of accuracy of the classification.

[1] See Adler, Martha, Mental Tests as a Basis for Classification, *Jour. of Educ. Psych.*, Vol. V, No. 1.

SECTION 9.—PRACTICAL APPLICATIONS IN HIGH SCHOOL CLASSIFICATION

There can be little question as to which of the three sources of estimate of a pupil's scholastic ability is the preferable one to use, in case it is not desired or possible to use all of them. The elementary school records of the pupils give the most accurate estimate of average class standing, as well as of standing in specific courses. A higher correlation than .80 between estimation and actual first-year standing should not be demanded, or expected, in a correlation of this nature.

There would be a great advantage in having a uniform record card, for each state school system, to contain, in addition to other data, the pupil's grades from year to year, together with a definite statement of the significance of the grades in terms of a normal distribution, or as deviations from the grade mean for the local system in question, expressed as multiples of the variability for that system. If these cards were freely transferred from school to school, as the pupil changed, it not only would be possible to classify pupils accurately each year, but it would be of incalculable value from other standpoints as well, for there is probably no easily obtainable data which could compare in significance with such a record.

The estimates of several of the previous teachers of the pupil give an excellent basis for classification, but wherever available, the more valuable records in the elementary school are probably also available, so that, for high school classification, they are not of prime importance.

They are, however, of specific value in analyzing the elements which contribute to scholastic success. In the regression equation based upon teachers' estimates, effort shows itself to be a very important factor. There would, therefore, be many advantages for educational, and even more particularly for vocational, guidance if there were available grades representing ability and effort, as well as accomplishment.

The importance of the interest and specific subject tests is not

to be measured solely by the extent of their correlation with class standing, as they probably are not at all measures of conscientiousness. Conscientiousness has been shown to be second in importance to intellectual ability only and to deserve a weight of 4 to 12 for all other factors measured, intellectual ability, emotional interest, and oral expression. A classification of pupils which does not take into account conscientiousness may be particularly advantageous in that it throws the indolent in with conscientious pupils of equal mentality, thus acting as a strong spur to the lazy while, at the same time, the group is homogeneous so far as capacity is concerned and it does not require a dual technique of presentation on the teacher's part to answer the needs of dull and bright pupils. It may be that in a small way a different technique of presentation is needed to best present a subject to a lazy pupil, from that needed in presenting it to an industrious one of the same mentality, but the difference does not compare with that needed in the case of dull and bright pupils.

It is also undoubtedly true that the tests here given, if given in a high school with classes one year apart, would yield higher correlations than here obtained. In the school from which groups 1, 2, and 3 came classification is close and grades differ by one-half year. In the school from which groups 4 and 5 came particular attention is paid to classification, resulting in courses which fit the needs of the pupil probably fully as thoroughly as in the other school. In fact, in both of these schools, certain classes differ from each other by not more than a quarter of a year. The effect of this is to make the classes more homogeneous and such homogeneity always decreases the correlation. Reference to the tables of means, p. 68, shows that the groups differ materially in their average accomplishments in the various tests. This shows that by means of these tests pupils could be classified as to their most probable place in the high school much more accurately than they can be classified as to their place in a class. The former is not the question which it is attempted to answer, but it is mentioned to show that the ability to place a pupil among all pupils is considerably greater than the ability to place him in a more or less homogeneous group, and as the groups here considered are unusually homogeneous the signif-

icance of the tests is correspondingly greater than the correlation coefficients indicate.

To make the classification still more reliable it is recommended that in the case of a pupil whose previous record is not available the tests here given be supplemented by acquired knowledge tests, particularly in mathematics and foreign languages.

SECTION 10.—GUIDANCE METHODS

It will be found that having once initiated a guidance bureau the demands upon it will be positive and innumerable—many of them extravagant. In the attempt to meet these demands, and to meet them on the spot and without a moment's delay, one of the richest sources of information is likely to be only very partially utilized. Reference is made to that product accumulated by every pupil—school grades. Whatever capacity it is that a grade, say, in mathematics, stands for, it is measured with a high degree of accuracy when the records of several years and of several teachers are combined. A pupil's school record is the most complete, detailed and accurate of all records, of the ordinary pupil, from his entrance in school to his entrance into work. Unless the significance of this record is evaluated with reference to all the important studies and vocations the most readily available and accurate data concerning the applicant for a place in some class, or for a job, are not being utilized. The evaluation of these data will require much statistical work, but its use after evaluation is simple.

Teachers' estimates of a pupil are second in importance only to grades. It requires, however, the estimates of several teachers to secure an accurate rating, and, under present conditions, it frequently is not possible to secure the estimate of more than one and, in cases where either the pupil or the principal changes location, even this is lacking. If each teacher were to place on record, at the end of every course, an estimate of several of the qualities, important for success, of each of his pupils, these data would be of inestimable value to the guidance expert. In this case, as that of school grades, a uniform record card, carrying a standardized grading, is essential for the best results.

The use of special tests in vocational guidance is unlimited. There could well be certain specialized tests for each important vocation, but first of all there might be devised a general test, somewhat along the line of the interest test in this study, which would have significance in all vocations and which could be evaluated with reference to any one desired. This general test

could test interest and general mental capacity, while it could be left to the specialized tests to measure specific capacity and the necessary acquired knowledge.

In so far as guidance becomes a science and not an intuition, in so far as its method and conclusions are capable of definition and free use by different individuals and are not simply inner convictions of the expert making them, the problem of relationship, expressed in quantitative terms, between the capacity of the applicant and the demands of the position will become more and more insistent for solution. A guidance bureau should be like a type distributing machine, which will take a hopperful of type, of all the letters of the alphabet, and place each in its particular niche, in the one place of all places where it fits. That a fitting distribution of human talent is a task of unmeasured intricacy is apparent, but the peculiar service thereby rendered to groping humanity makes the solution worthy the greatest effort.

In broad outline, as already pointed out, the problem of vocational guidance consists of measuring the demands of the possible vocations, and of the capacities of the applicant and then fitting the applicant into that place which best suits his talents and his ambitions. In detailed procedure, the regression equation method is a powerful instrument, for it enables any number of factors to be combined with the highest significance with reference to the vocation in question. When a large number of factors, none of them of predominant importance, contribute to a total result, the human intellect, unaided, cannot compass their total significance and it is only by mathematical means that they can be summed and interpreted.

SECTION 11.—APPENDIX

Ages of Pupils

This study covers four different groups of pupils: (1) 59 pupils starting the second year of the high school of School A; (2) 42 starting the second term of the first year of the same school; (3) 81 starting the first term of the first year of the same school; (4) 26 pupils starting the second year of the high school of School B, and (5) 25 pupils starting the first year of the high school of School B. Ages are expressed in years and tenths of a year from birth up to January 1, 1913. Since the algebra and geometry tests were given during the last of September and the first of October, 1913, and the English, history and interest tests were given in January, 1913, the average ages at the times of the tests may be obtained from the given means by subtracting .30 of a year in the cases of the algebra and geometry tests and by adding .05 of a year in the cases of the other tests. The mean ages of the different groups January 1, 1913, is as follows:

Group 1	16.1 years.
Group 2	14.6 "
Group 3	13.7 "
Group 4	15.7 "
Group 5	14.4 "

The Assignment of Numerical Magnitudes for Literal Grades

In both schools a literal grading system is in use. In School A letters A, B, C, D and E are used. The mark E is used very infrequently—some teachers not using it at all. In averaging the grades for two or more terms it was assumed that the difference in ability represented by grades of A and B was equal to the difference in ability represented by grades of B and C, etc. That little error resulted from this assumption will be shown in the next section of this appendix. In averaging the grades of four terms the following differences in ability may occur:

$$A = \frac{A+A+A+A}{4},\quad A\tfrac{1}{2}- = \frac{A+A+A+B}{4},$$

$$B+ = \frac{A+A+B+B}{4} = \frac{A+A+A+C}{4},\quad B\tfrac{1}{2}+ = \frac{A+B+B+B}{4} = \text{etc.},$$

$$B = \frac{B+B+B+B}{4} = \frac{A+B+B+C}{4} = \text{etc.}$$

And so on for other combinations. The literal grades thus obtained were then transformed into numerical grades, assuming a normal distribution of talent. This is very readily done by noting the percentage frequencies of the different

grades and using such a table for transformation as that given by Thorndike in his "Mental and Social Measurements," pp. 221–225, second edition. Upon this basis literal grades were assigned numerical values as follows:

COURSES IN WHICH TESTS WERE GIVEN AT BEGINNING OF TERM

GROUPS	ALG. 1 yr. 2 and 3	ALG. 1 YR. 5	GEO. 1 YR. 1	GEO. 1 YR. 4	ENG. ½ YR. 1	ENG. ½ YR. 2	ENG. ½ YR. 3	ENG. ½ YR. 4 and 5	HIST. 3RD & 4TH YR. ½ YR. 1	HIST. 2ND YR. ½ YR. 1	HIST. 1ST. YR. ½ YR. 2
A+			2.44		1.99						
A½+											
A	2.34		1.42		1.49	2.70	1.81		2.23	1.23	1.96
A½−	1.69		.99								
B+	1.29		.88	1.56	1.27	1.65	1.13	1.49		.55	1.14
B½+	1.05	1.92	.77								
B	.80	.96	.51	1.28	.55	.78	.68	.67	1.29	.29	.67
B½−	.57	.44	.23	.92							
C+	.38	.06	.08	.58	.07	.15	.21	− .05	.77	− .05	.28
C½+	.17	− .24	− .08	.15							
C	− .01	− .58	− .20	− .36	− .48	− .28	− .39	− .70	.32	− .60	− .20
C½−	− .24	−1.20	− .31	− .71				−1.35			
D+	− .47	−2.10	− .53	−1.00	− .95	− .73	−1.13	−1.99	− .32	−1.32	− .74
D½+	− .64		− .81	−1.42							
D	− .77		−1.09	−2.16	−1.46	−1.11	−1.99		− .88	−2.16	−1.33
D½−			−1.49								
D−	−1.03		−1.89		−1.85	−1.34			−1.17		
E+	−1.54		−2.44		−2.10	−1.65					−1.86
E½+											
E	−2.34				−2.70	−2.28			−1.49		−2.62
E−									−2.23		

The same kind of transformation tables were obtained in 18 other courses in order to obtain numerical measures for the literal grades received in the first year of the high school by the 59 pupils whose records were available down to the third grade. The populations upon which the transformations were based averaged 40.4 pupils per course.

The grades of A+ and A½+ require explanation. In mathematics and English special classes were formed for the particularly bright pupils. The grading of pupils in these classes was more severe than the ordinary grading. It was the opinion of the teachers concerned that grades would be comparable with the rest of the grades of the school if the grades received in the special mathematics classes were raised one point, i.e., call C's, B's and B's, A's, etc., and if the grades received in the special English classes were raised one-half of a point. This was accordingly done, and accounts for the grades A+ and A½+.

The grade D− is an average of such grades as the following: first term D, second term E, third term E, fourth term D. This is a passing grade for the year. The grade E+ is an average of the same grades, except that the final term is an E, constituting a failure for the year. It is reasonable to assume that slightly greater proficiency is shown in the former case than in the latter.

A further simple transformation was made in order to obtain measures that were convenient to work with. The numerical measures obtained by use of the preceding transformation tables were each divided by .2 and the results kept to the nearest integer. The range thus obtained has about 26 divisions

in it and the standard deviation is about 5. This distribution is very convenient for purposes of calculation and the effect of the grouping is so slight that no correction in the value of the coefficients of correlation need be made on account of it. The distributions thus obtained have means at zero, to a very close approximation, and no correction to the coefficients of correlation is necessary to correct for arbitrary means.

To test the extent of the error due to the averaging of literal grades, the following facts are to be considered:

Extent of Error in Averaging Literal Grades

In averaging literal grades such as $\frac{\text{A plus B}}{2}=\text{B+}$; $\frac{\text{B plus C}}{2}=\text{C+}$, etc. no error is introduced because the only assumption involved is that A > B+ > B, etc., which is the basic assumption underlying the transformation of literal grades into numerical grades. Some 92 per cent of the averaging was of this nature. It is only when it is stated that $\frac{\text{A plus C}}{2}=\frac{\text{B plus B}}{2}=\text{B}$ that there is danger of error from this source. Simplifying we find that this equation is true only in case A − B = B − C. The following data show to how close an extent this assumption is true and it should be remembered that it applies to only about 8 per cent of the averaging done.

	Algebra-Groups 2 and 3. Average of Two Quarters	English-Groups 1, 2 and 3. Average of Two Quarters	History-Group 1. Average of Two Quarters	Geometry-Group 1. Average of Two Quarters
	Corresponding grade:	Corresponding grade:	Corresponding grade:	Corresponding grade:
A	1.41	1.80	1.57	2.03
	A − B = .86	A − B = 1.04	A − B = 1.05	A − B = 1.22
B	.55	.76	.52	.81
	B − C = .71	B − C = 1.00	B − C = .80	B − C = .91
C	− .16	− .24	− .28	− .10
	C − D = .86	C − D = 1.10	C − D = .99	C − D = .89
D	−1.02	−1.34	−1.27	− .99
	D − E = .93	D − E = 1.27	D − E = 1.06	D − E = 1.04
E	−1.95	−2.61	−2.33	−2.03
	E − F = .88			
F	−2.83			

From the above table:

	If A − B = 1, then B − C =	If B − C = 1, then C − D =	If C − D = 1, then D − E =	If D − E = 1, then E − F =
Alg.	.83	1.21	1.08	.91
Eng.	.96	1.10	1.15	
Hist.	.76	1.24	1.07	
Geom.	.75	.98	1.17	
Av.	.83	1.13	1.12	

Average of all = 1.02

From similar tables for groups 4 and 5:

	If A−B=1, then B−C=	If B−C=1, then C−C−=	If C−C−=1, then C−−D=	
Alg.	.842	.679	.99	
Eng.....	.914	1.018		
Geom.	1.070	.855	.93	
Av.....	.942	.851	.96	

Average of all = .918

Similar data from the elementary school group show a still closer approach to equality.

It is therefore plain that no appreciable error has been introduced by such averaging.

Elementary School Grades

In the elementary school the system of grading for certain years was different from that for other years. In a few of the grades the literal system A, B, C, D, E, F, was used, but in the major number of grades considered the marks given were 1, 2, 3,—1 being the highest grade used. By assuming a normal distribution, and expressing both the values 1, 2, 3 and A, B, C, D, E, F in terms of deviations from the means, the values may be compared with each other. The following relation exists:

English	Arithmetic	History	Average Weighted Relation (Used in all transformations.)
A to F 1 to 3			
A= 1.2σ=1.0	A= 1.8σ= .9	A= 1.7σ= .9	A= .9
B= .2σ=1.6	B= .8σ=1.6	B= .5σ=1.8	B=1.6
C=− .6σ=2.1	C= .0σ=2.1	C=− .5σ=2.2	C=2.1
D=−1.3σ=2.5	D=− .8σ=2.6	D=−1.4σ=2.8	D=2.6
E=−1.7σ=2.7	E=−1.6σ=3.1	E=−2.2σ=3.3	E=3.0
F=−2.4σ=3.1	F=−2.4σ=3.6		F=3.4

With this transformation table literal grades were expressed in units that are comparable with the numerical grading 1, 2, 3. The final grade given each year, for each pupil, in English is the sum of the grades given for the four terms of the school year in two English courses, e.g. "Reading and Literature" and "Composition." Thus the poorest grade possible is 24 and the best 8. The grades given in arithmetic and history are the sum of the grades for the four terms for these subjects, therefore the lowest grade possible is 12 and the highest 4 in each of them.

A few of the 59 pupils in this group skipped a year. In such a case the grade of the year before was entered for the year skipped. This, of course, would be a high grade and representative of the ability of the pupil.

In expressing the grades of these 59 pupils as deviations from the means, the means of the grades in question were obtained. Random samplings of about 40 from each of the grades for the various years were the basis for the calculation of the means. The reason for this is apparent. Since this is a prognosis problem, the prognosis must be based upon the rankings of the individuals in the groups in which they are found.

Several investigations have shown that there is a selective process operating to eliminate backward pupils from the grammar grades. Under these conditions, the 7th grade average of those who attend the high school would be slightly greater than the average of all 7th grade pupils, still more greatly above the average of all 6th grade pupils, etc. The difference would be most pronounced when dealing with the average grade of these pupils in the 4th grade. Calculation shows that this particular group of 59 pupils is .224 sigma above the average of first year pupils in general. Calculation also shows that they are .367 sigma above the average of 4th grade pupils. There is thus only the small difference of .143 sigma which can be attributed to selective elimination of the weaker pupils. Taking all the elementary grades together, as combined by the regression equation, it is found that there is not even a difference of .143 sigma. In fact, these pupils who, as first year pupils, are .224 sigma above the average of such pupils are, as elementary pupils, only .186 sigma above the average of 7th to 4th grade pupils when grades are weighted according to the regression equation. If any conclusion is justified from this small population, it is that in this particular school there is no selective elimination of the duller pupils.

In calculating the correlation between first year standing and the combined 7th to 4th grade standing, no correction is made due to the means of the population of 59 being different from the means of the entire body of first-year high school pupils, or the entire body of 7th to 4th grade pupils. Due to the nature of the problem no correction to the first-year high school mean is permissible as the object of the study is to prognosticate divergence from this mean. A correction to the mean of the combined elementary grades might be applied but its magnitude would be .224 sigma $-.186$ sigma $= .038$ sigma, which is negligible. Even the correction to the 4th grade alone, .224 sigma $-.367$ sigma $= -.143$ sigma, is inconsequential.

The calculation of the regression equation giving the regression of the first-year high school grades upon a combination of 7th to 4th grade marks is based upon the following data:

	F_A	7_A	6_A	5_A	4_A
7_A	.719				
6_A	.728	.730			
5_A	.531	.425	.541		
4_A	.624	.551	.573	.576	
σ's	3.796	4.167	4.891	5.308	5.912

$$F_A = .3540\frac{2.250}{2.585}(7_A) + .3113\frac{2.250}{2.911}(6_A) + .1352\frac{2.250}{4.080}(5_A) + .2419\frac{2.250}{4.215}(4_A)$$

$$F_A = .3082\,(7_A) + .2407\,(6_A) + .0746\,(5_A) + .1291\,(4_A)$$

or approximately $1.803\ (F_A) = \frac{1}{3}[1.667\ (7_A) + 1.3\ (6_A) + .4\ (5_A) + .7\ (4_A)]$. This combination of elementary school records is designated as $(7, 6, 5, 4_A)$, and the correlation, $r_{F_A(7, 6, 5, 4_A)} = .789$.

The same equation is used to obtain measures in elementary school mathematics and English, except that division by three is omitted, giving the following:

$$(7, 6, 5, 4_M) = 1.667\ (7_M) + 1.3\ (6_M) + .4\ (5_M) + .7\ (4_M)$$
$$(7, 6, 5, 4_E) = 1.667\ (7_E) + 1.3\ (6_E) + .4\ (5_E) + .7\ (4_E)$$

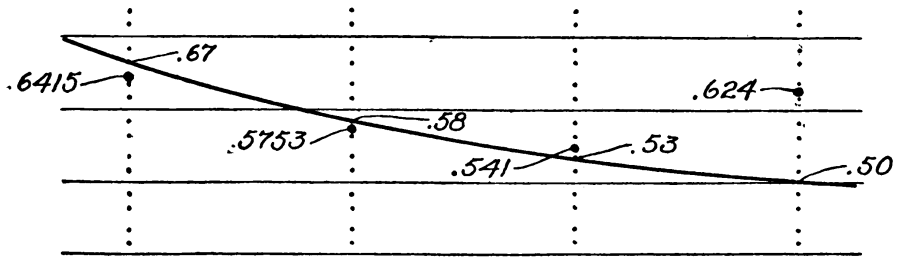
Calculation gives:

$$r_{F_M(7, 6, 5, 4_M)} = .580 \text{ and } r_{F_E(7, 6, 5, 4_E)} = .710$$

No history was taken during the first high school year so there are no history correlations.

It may be noticed by reference to the preceding table that $r_{F_A 7_A}$ and $r_{F_A 5_A}$ are less than would be expected from the other correlation coefficients. This may be due to the teachers of these particular 7th and 5th grades being less expert in estimating the ability of pupils than the 6th and 4th grade teachers. Whatever the cause, probably a better regression equation for general purposes can be obtained than the one given above. The accompanying curve

r Between Grade in a Given Year and Grade One Yr. Before	*r* Between Grade in a Given Year and Grade Two Yrs. Before	*r* Between Grade in a Given Year and Grade Three Yrs. Before	*r* Between Grade in a Given Year and Grade Four Yrs. Before
Av. of 4 coef's. = .6415	Av. of 3 coef's. = .5753	Av. of 2 coef's. = .541	1 coef. = .624



was drawn with this end in view. A smooth curve, not rectilinear, is drawn near the points representing the ordinates for the various abscissae. The intersections of the curve with the ordinates give the values of the correlation coefficients in the succeeding table.

The falling off in correlation from year to year is thought to be reasonable and calculation will show that the sum of the deviations of the actual coefficients of correlation from the points where the curve crosses the ordinates at the various abscissae very nearly equals zero, so that the curve is not entirely arbitrary.

	Yr. in Question 0	1 Yr. Before 1	2 Yrs. Before 2	3 Yrs. Before 3	4 Yrs. Before 4
1	.67				
2	.58	.67			
3	.53	.58	.67		
4	.50	.53	.58	.67	
σ's	σ_0	σ_1	σ_2	σ_3	σ_4

The following regression equation is derived from the table:

$$X_0 = .4616\frac{\sigma_0}{\sigma_1}X_1 + .1458\frac{\sigma_0}{\sigma_2}X_2 + .0910\frac{\sigma_0}{\sigma_3}X_3 + .1098\frac{\sigma_0}{\sigma_4}X_4$$

In case the standard deviations are all equal, this equation becomes, to a very close approximation: $54.9\,X_0 = 25\,X_1 + 8\,X_2 + 5\,X_3 + 6\,X_4$.

Teachers' Estimates and Combinations of the Same

Having the estimates of several teachers of the mental traits of each pupil, it is necessary to combine these estimates into single measures of the trait in question for the pupil in question. Each teacher reported on about 25 pupils, and it is assumed that the talent follows a normal distribution. The rankings of the teachers were then expressed as deviations from their respective means, using the same method as used in transforming literal into numerical grades. The number of gradings obtained for each pupil ranges from two to seven. It will be seen from the following table that the correlation between these estimates is low:

$$r_{(\text{i. a. —first judge})(\text{i.a.—second judge})} = .28 \pm .024$$
$$r_{(\text{cons. } \text{“} \text{ })(\text{cons. } \text{“} \text{ })} = .38 \pm .022$$
$$r_{(\text{emo. int. } \text{“} \text{ })(\text{emo. int. } \text{“} \text{ })} = .31 \pm .024$$
$$r_{(\text{exp. } \text{“} \text{ })(\text{exp. } \text{“} \text{ })} = .29 \pm .024$$

Because of this, it is impossible to average these grades and have them even approximately comparable, for the standard deviation of the measures which are the averages of the grades of two judges is materially greater than the standard deviation of the measures which are the averages of the grades of a greater number of judges. The extent of this difference can be readily estimated, for if X_1, X_2, $\cdots$ X_n are measures of the same trait, and if r is equal to the correlation between such measures, and if sigma X_1 = sigma X_2 $= \cdots$ sigma X_n, then we have:

$$\sigma_{\frac{X_1+X_2}{2}} = \sqrt{\frac{\Sigma(X_1+X_2)^2}{4n}} = \sqrt{\frac{\Sigma(X_1^2+2X_1X_2+X_2^2)}{4n}}$$
$$= \sqrt{\frac{2\Sigma X_1^2 + 2nr\sigma_{X_1}\sigma_{X_2}}{4n}} = \sigma_{X_1}\sqrt{\frac{1+r}{2}}$$

Similarly,

$$\sigma_{\frac{X_1+X_2+X_3}{3}} = \sigma_{X_1}\sqrt{\frac{1+2r}{3}}$$

Finally,

$$\sigma_{\frac{X_1+X_2+\cdots X_n}{n}} = \sigma_{X_1}\sqrt{\frac{1+(n-1)r}{n}}$$

In order to make the standard deviations of measures which are averages of varying numbers of estimates comparable, it is only necessary to divide the measures by their respective sigmas, i.e., if but a single measure divide by sigma X_1; if an average of two measures divide by sigma $\frac{X_1+X_2}{2}$, etc. The following table gives the desired divisors for the various cases:[1]

No. of grades averaged	I. a. Intellectual ability	Cons. Conscientiousness	Emo. i. Emotional interest	Exp. Expression
1	.92	.92	.92	.95
2	.74	.79	.74	.76
3	.66	.73	.66	.68
4	.63	.69	.63	.65
5	.61	.67	.61	.63
6	.59	.66	.59	.61
7	.58	.65	.58	.60

The mean number of grades averaged to obtain each individual's measure is about two and one-half, therefore the reliability coefficients have approximately the following values:

$r_{\text{(i. a. measures as derived) (i. a. meas. similarly derived)}} = .493$
$r_{\text{(cons. " ") (cons. " " ")}} = .605$
$r_{\text{(emo. i. " ") (emo. i. " " ")}} = .505$
$r_{\text{(exp. " ") (exp. " " ")}} = .529$

In 23 cases of group 2, teachers' estimates were not available, except teachers of English and history who later had the same pupils in test courses, which were continuation courses of the first half year's work under the same teacher. The estimates of such teachers were not used when it could be avoided, i.e., in all cases except these 23. The following correlations justify excluding such estimates:

$r_{\text{G(i. a.—estimate of geometry teachers)}} = .57$ { pop. = 57
$r_{\text{G(i. a.— " " other ")}} = .44$ { group 1

$r_{\text{E(i. a.— " " English ")}} = .54$ { pop. = 33
$r_{\text{E(i. a.— " " other ")}} = .36$ { group 2

[1] This method of averaging varying numbers of correlated measures was used frequently in other portions of this study, e.g., in averaging grades of pupils for some given term or year where the number of studies varied appreciably.

The estimates of "other" teachers may be considered as accurate as those of the geometry or English teachers, so that the excess of .57 over .44, and of .54 over .36 is, in a sense, a measure of the extent to which a teacher's estimate is based upon the unique ability shown in the subject he teaches.

The combination of the measures, based upon teachers' estimates, into a single final measure or estimate of scholastic ability is accomplished by the usual regression equation, as follows:

	Av.	I. a.	Cons.	Emo. i.	Exp.
I. a.	.72				
Cons.	.62	.61			
Emo. i.	.58	.61	.66		
Exp.	.63	.82	.55	.59	
σ's	4.048	5.193	5.166	5.138	5.190

$$\text{Av.} = .3584\frac{2.630}{2.591}(\text{I. a.}) + .2456\frac{2.630}{3.530}(\text{Cons.}) + .1161\frac{2.630}{3.547}(\text{Emo. i.}) + .0471\frac{2.630}{2.905}(\text{Exp.})$$

or Av. = .364(I. a.) + .183 (Cons.) + .086 (Emo. i.) + .043 (Exp.)
or, approximately, 1.1 (Av.) = .4 (I. a.) + .2 (Cons.) + .1 (Emo. i.) + .05 (Exp.)
—which is a very simple equation to use.

$$r_{\text{(Av.) (I. a., Cons., Emo. i., Exp.)}} = .7551.$$

Bearing of the Various Factors, I. a., Cons., Emo. i., and Exp., upon M, E, and H.

From the accompanying data:

$$M_{t.\,e.} = .460\frac{\sigma_M}{\sigma_{I.\,a.}}\text{I. a.} + .114\frac{\sigma_M}{\sigma_{Cons.}}\text{Cons.} + .129\frac{\sigma_M}{\sigma_{Emo.\,i.}}\text{Emo. i.} - .014\frac{\sigma_M}{\sigma_{Exp.}}\text{Exp.}$$

	M	I. a.	Cons.	Emo. i.	Exp.
I. a.	.591				
Cons.	.467	.61			
Emo. i.	.472	.61	.66		
Exp.	.496	.82	.55	.59	
σ's	σ_M	$\sigma_{I.\,a.}$	$\sigma_{Cons.}$	$\sigma_{Emo.\,i.}$	$\sigma_{Exp.}$

$r_{MM_{t.\,e.}} = .61$ (Population = 178).

From the accompanying data:

$$E_{t.\,e.} = .336\frac{\sigma_E}{\sigma_{I.\,a.}}\text{I. a.} + .251\frac{\sigma_E}{\sigma_{Cons.}}\text{Cons.} + .068\frac{\sigma_E}{\sigma_{Emo.\,i.}}\text{Emo. i.} + .083\frac{\sigma_E}{\sigma_{Exp.}}\text{Exp.}$$

	E	I. a.	Cons.	Emo. i.	Exp.
I. a.	.598				
Cons.	.546	.61			
Emo. i.	.487	.61	.66		
Exp.	.536	.82	.55	.59	
σ's	σ_E	$\sigma_{I.\,a.}$	$\sigma_{Cons.}$	$\sigma_{Emo.\,i.}$	$\sigma_{Exp.}$

$r_{E\,E_{t.\,e.}} = .64$ (Population = 179)

From the accompanying data:

$$H_{t.e.} = .450\frac{\sigma_H}{\sigma_{I.a.}} I.a. - .024\frac{\sigma_H}{\sigma_{Cons.}} Cons. + .305\frac{\sigma_H}{\sigma_{Emo.i.}} Emo.i. - .287\frac{\sigma_H}{\sigma_{Exp.}} Exp.$$

$$r_{H H_{t.e.}} = .46 \quad (Population = 68)$$

	H	I. a.	Cons.	Emo. i.	Exp.
I. a.	.381				
Cons.	.290	.61			
Emo. i.	.390	.61	.66		
Exp.	.245	.82	.55	.59	
σ's	σ_H	$\sigma_{I.a.}$	$\sigma_{Cons.}$	$\sigma_{Emo.i.}$	$\sigma_{Exp.}$

Grading of the Algebra Test

Having the gradings for the various problems in the algebra test, it is impossible to say, *a priori*, with any assurance, which are the most significant and which are the least so. The common procedure in a case like this is to call them all of equal importance and add or average. Whether such a procedure results in getting out of the data all that is in them or not, is a fit subject for investigation. The question is simply this—does the magnitude (grade of prob. 1+ grade of prob. 2 $\cdots$ + grade of prob. 14) correllate as highly with the algebra grade received at the end of the school year, as the magnitude ($C_1 \times$grade of prob. $1 + C_2 \times$grade of prob. $2 + \cdots C_{14} \times$grade of prob. 14) where $C_1, C_2, \cdots C_{14}$ have the best values possible. Of course the second magnitude would result in the higher correlation, or, what amounts to the same thing, the standard deviation of the residuals in the second case is smaller than the standard deviation of the residuals in the former case. Using the notation given by Yule, this is to say that

$$\sigma_{A.\,1,2,3,4,5,6,7,8,9,10,11,12,13,14} < \sigma_{A.\,1+2+3+4+5+6+7+8+9+10+11+12+13+14}$$

(A = grade in algebra course, 1 = grade of first problem, etc.) It is manifestly impractical to attempt to calculate $\sigma_{A.\,1,2,3,4,5,6,7,8,9,10,11,12,13,14,}$ but an approximation to this may be obtained if the problems 1, 2, $\cdots$ 14 that have about the same standard deviations, and that are correlated to about the same extent with A, and are correlated with each other to approximately an equal extent, are grouped, thus reducing the variables to such a number that the calculation is feasible. In attempting to fulfill these conditions, problems 1–4, 8–11, were grouped, as were also problems 6, 13, 14, giving groups A', B', C', respectively. The question then is to determine $\sigma_{A.\,A'B'C'}$ and $\sigma_{A.\,A'+B'+C'}$. Formulae giving these expressions (derived in the next section) are as follows:

$$\sigma^2_{A.\,A'B'C'} = \sigma^2_A\left[1 - \frac{\Sigma r^2_{AA'} - 2\Sigma r_{AA'}r_{AB'}r_{A'B'} - \Sigma r^2_{AA'}r^2_{B'C'} + 2\Sigma r_{AA'}r_{AB'}r_{A'C'}r_{B'C'}}{1 - \Sigma r^2_{A'B'} + 2\, r_{A'B'}r_{A'C'}r_{B'C'}}\right]$$

$$\sigma^2_{A.\,A+B'+C'} = \sigma^2_A\left[1 - \frac{(\Sigma r_{AA'}\sigma_{A'})^2}{\Sigma\sigma^2_{A'} + 2\,\Sigma r_{A'B'}\sigma_{A'}\sigma_{B'}}\right]$$

The various coefficients of correlation and standard deviations required, are given in the following table:

	A	A'	B'	C'
A'	.39			
B'	.41	.17		
C'	.39	.41	.42	
σ's Estimated		10.3	7.0	7.0

From which
$$\sigma^2_{A.\,A'B'C'} = .705\,\sigma^2_A.$$
$$\sigma^2_{A.\,A'+B'+C'} = .711\,\sigma^2_A.$$

These coefficients were calculated by the per cent of unlike signs method and only approximate to Pearson coefficients of correlation, which are necessary from the theoretical standpoint, but it is thought that they satisfactorily serve the purpose of this preliminary investigation.

The difference between the standard deviations given above is so small that the advantage of the regression method over the simple average or sum method is plainly too slight to justify the added work, and the measure for the algebra test is taken simply as the sum of the grades in the test, after subtracting the mean and dividing, for convenience, by 5.

That is,
$$A_t = \frac{A'+B'+C'-\text{mean}}{5}$$

Here, as in the case with all measures used in regression equations, A_t is a deviation from the mean.

Derivation of Formulae

$$\sigma^2_{1\cdot 2} = \sigma^2_1(1-r^2_{12})$$ Formula for 2 variables.

$$\sigma^2_{1\cdot 23} = \sigma^2_1(1-r^2_{12})(1-r^2_{13\cdot 2})$$
$$= \sigma^2_1\left[1-\frac{r^2_{12}+r^2_{13}-2\,r_{12}r_{13}r_{23}}{1-r^2_{23}}\right]$$

or
$$= \sigma^2_1\left[1-\frac{\Sigma r^2_{12}-2\,\Sigma r_{12}r_{13}r_{23}}{1-\Sigma r^2_{23}}\right]$$ Formula for 3 variables.

$$\sigma^2_{1\cdot 234} = \sigma^2_1(1-r^2_{12})(1-r^2_{13\cdot 2})(1-r^2_{14\cdot 23})$$
$$= \sigma^2_1(1-r^2_{12})\left(1-\frac{1-r^2_{13\cdot 2}-r^2_{34\cdot 2}r^2_{14\cdot 2}+2\,r_{13\cdot 2}r_{14\cdot 2}r_{34\cdot 2}}{1-r^2_{34\cdot 2}}\right)$$
$$= \sigma^2_1\left[1-\frac{\Sigma r^2_{12}-2\,\Sigma r_{12}r_{13}r_{23}-\Sigma r^2_{12}r^2_{34}+2\,\Sigma r_{12}r_{13}r_{24}r_{34}}{1-\Sigma r^2_{23}+2\,r_{23}r_{24}r_{34}}\right]$$ Formula for 4 variables.

where $\Sigma r_{12}r_{13}r_{23} = r_{12}r_{13}r_{23} + r_{12}r_{14}r_{24} + r_{13}r_{14}r_{34}$ $\Sigma r^2_{12} = r^2_{12} + r^2_{13} + r^2_{14}$

$\Sigma r^2_{12}r^2_{34} = r^2_{12}r^2_{34} + r^2_{13}r^2_{24} + r^2_{14}r^2_{23}$

$\Sigma r_{12}r_{13}r_{24}r_{34} = r_{12}r_{13}r_{24}r_{34} + r_{12}r_{14}r_{23}r_{34} + r_{13}r_{14}r_{23}r_{24}$

$\Sigma r^2_{23} = r^2_{23} + r^2_{24} + r^2_{34}$

$$\sigma^2_{1.2+3} = \sigma^2_1(1 - r^2_{1.2+3}) = \sigma^2_1\left[1 - \frac{(r_{12}\sigma_1\sigma_2 + r_{13}\sigma_1\sigma_3)^2}{\sigma^2_1(\sigma^2_2 + \sigma^2_3 + 2r_{23}\sigma_2\sigma_3)}\right]$$

$$= \sigma^2_1\left[1 - \frac{(r_{12}\sigma_2 + r_{13}\sigma_3)^2}{\sigma^2_2 + \sigma^2_3 + 2r_{23}\sigma_2\sigma_3}\right]$$

Similarly,

$$\sigma^2_{1.2+3+4} = \sigma^2_1\left[1 - \frac{(r_{12}\sigma_2 + r_{13}\sigma_3 + r_{14}\sigma_4)^2}{\sigma^2_2 + \sigma^2_3 + \sigma^2_4 + 2(r_{23}\sigma_2\sigma_3 + r_{24}\sigma_2\sigma_4 + r_{34}\sigma_3\sigma_4)}\right]$$

$$= \sigma^2_1\left[1 - \frac{(\Sigma r_{12}\sigma_2)^2}{\Sigma\sigma^2_2 + 2\Sigma r_{23}\sigma_2\sigma_3}\right]$$

It may be easily shown that this last formula is general and holds good for any number of variables.

It was attempted to derive a formula for $\sigma_{1.2345}$, but paper and patience were exhausted before it was accomplished. The law governing the coefficents of the terms of the formula for $\sigma_{1.234}$ is not sufficiently clear to enable the author to state a general formula applicable to residuals of higher order.

Grading of Geometry Test

The plan adopted in connection with the algebra test is used in combining the parts of the geometry test. The following table of coefficients of correlation (calculated by per cent of unlike signs method) supplies the necessary data:

	G	E'	F'	G'	H'
E'	.25				
F'	−.06	.13			
G'	.22	.19	.06		
H'	.40	.13	.09	.45	
Estimated σ's		3.3	7.4	4.0	9.0

E' = problem 1; F' = problems 2, 4, 5, 6; G' = problem 7; H' = problems 8, 9, 10. Problem 3 turned out to be too easy for the group in question, some 98 per cent making a perfect score, so its grading is not used. The surprise in this table is the small negative correlation between G and F'. Because of this small correlation F' is also discarded. For the balance of the data the advantage of the regression method over the average method is negligible, as the following standard deviations show:

$\sigma^2_{G.\ E'G'H'} = .800\,\sigma^2_G$ $\qquad$ $\sigma_{G.\ E'G'H'} = .894\,\sigma_G$

$\sigma^2_{G.\ E'+G'+H'} = .817\,\sigma^2_G$ $\qquad$ $\sigma_{G.\ E'+G'+H'} = .904\,\sigma_G$

$\sigma_{G.\ E'+F'+G'+H'} = .925\,\sigma_G$

Accordingly the sole measure of the geometry test is taken as the average of E', G', H'; $=$ i.e. $G_t \frac{E'+G'+H'-\text{mean.}}{3}$

Grading of the English Test

The following gradings of the English and history tests may be of particular significance as evidence of ability in English: E_a (accuracy of description in the English test), E_v (valuation of important factors in the English test), W (written expression in both English and history tests), D (dramatization in both English and history tests).

The correlations between dramatization and English and history were calculated by the percentage of unlike signs method, to determine with which of these subjects dramatization is most closely associated, with the result that $r_{ED} = .40$ (later calculation gave Pearson coefficient of correlation to equal .182) and $r_{HD} = .09$. The correlations between English and the other elements, mentioned above, are as follows: $r_{EE_a} = .64$, $r_{EE_v} = .59$, $r_{EW} = .56$, when calculated by the unlike signs method. (Later calculations show that these are somewhat larger than the Pearson method would give.) Furthermore, the correlation between written expression and dramatization is .61 (by the unlike signs method). The grading for dramatization is seen to be more closely related to English than history, but because of its high correlation with written expression it contributes little that is unique, even with reference to English and is, accordingly, not evaluated with respect to either English or history.

Another factor in connection with dramatization is the question of whether its relation to English or history is linear. In grading the papers, the author was quite impressed with the feeling that high grades in dramatization were more likely to accompany good or poor grades in English than medium grades. The author was totally unaware of the English grades of the pupils at the time of the grading for dramatization, so there was no reliable foundation for the belief. After the grades were available this question was tested by inspection of the regression line in the correlation table for English and dramatization and by calculation of the correlation ratio. The regression line was irregular, but did show some evidence of such non-rectilinearity. The value of the correlation ratio between English and dramatization is .241. (Compare with $r_{ED} = .182$.) The excess of .241 over .182 is not sufficient to warrant the assertion that the regression is not rectilinear, according to the criterion established by Blakeman,[1] but the chances are in favor of its being non-rectilinear, so the question is still an open one.

To return to the determination of the method of combining the parts of the English test: the following table of coefficients of correlation (unlike signs method) gives the necessary data:

$\sigma^2_{E.\ E_aE_vW} = \sigma^2_{E.569}$

$\sigma^2_{E.\ E_a+E_v+W} = \sigma^2_{E.571}$

	E	E_a	E_v	W
E_a	.64			
E_v	.59	.86		
W	.56	.75	.64	
Estimated σ's		2.5	2.2	1.65

The advantage of the regression equation method is so small as not to justify the added labor necessitated by its use, and the simple average, for convenience multiplied by 2, of E_a, E_v and W, is taken as the measure in the English test. That is, $E_t = \frac{2}{3}(E_a + E_v + W - \text{mean})$.

[1] See J. Blakeman, Biometrika Vol. 4, pp. 349–50, for criterion of rectilinearity: Here the function of η and r in question = 1.69, which is less than 2.5 the required value if non-rectilinearity is to be definitely established.

Grading of History Test

The population in the case of history is small and therefore Pearson coefficients of correlation were calculated for the preliminary investigation, instead of the less accurate coefficients calculated by the percentage of unlike signs method. The data are as follows:

$\sigma^2_{H.\ H_aH_v} = .898\,\sigma^2_{H.}$

$\sigma^2_{H.\ H_a+H_v} = .906\,\sigma^2_{H.}$

$\sigma^2_{H.\ H_a} = .904\,\sigma^2_{H.}$

	H	H_a	H_v
H_a	.31		
H_v	.23	.62	
σ's		2.40	1.69

$r_{HH_a} = \sqrt{.096} = .310$, and similarly we may say that the total correlation between history and the combined measures H_a, H_v is given by the following expressions: $r_{H(H_a+H_v)} = \sqrt{.094} = .307$ and $r_{H(H_aH_v)} = \sqrt{.102} = .319$, where the notation $r_{H(H_aH_v)}$ is understood to mean the correlation between history and H_a and H_v when combined into a single measure by the regression equation. The above results show that the average, or sum, H_a+H_v, will give a lower correlation than H_a alone, and that the regression equation yields but .009 higher correlation. For these reasons the sole measure of the history test is taken to be H_a, for convenience multiplied by two. That is, $H_t = 2(H_a - \text{mean})$.

Bearing of the Various Tests Upon Mathematics

(a) *Algebra.*

To evaluate the significance of the algebra, English and history tests in their bearing upon algebra, the regression equation between these tests and algebra grades may be calculated. The following table gives the required data:

$\sigma^2_{A.\ A_tE_tH_t} = .728\,\sigma^2_A$

$\sigma^2_{A.\ A_t+E_t+H_t} = .776\,\sigma^2_A$

	A	A test	E test	H test
A_t	.47			
E_t	.37	.37		
H_t	.27	.27	.40	
σ's	4.977	3.856	3.286	5.460

There is here a material advantage to be gained by the use of the regression method, and it has accordingly been calculated:

$$A_{ct} = 1.316\,(A) = .6\,(A_t) + .4\,(E_t) + .11\,(H_t)$$

(b) *Geometry.*

A procedure, similar to the above for algebra, gives the following results:

$\sigma^2_{G.\ G_tE_tH_t} = .820\,\sigma^2_G$

$\sigma^2_{G.\ G_t+E_t+H_t} = .856\,\sigma^2_G$

	G	G_t	E_t	H_t
G_t	.42			
E_t	.24	.42		
H_t	.21	.20	.40	
σ's	5.010	3.751	3.042	5.176

$G_{ct} = 1.53(G) = .8(G_t) + .08(E_t) + .183(H_t)$ or $G_{ct} = .8[(G_t) + .1(E_t) + .23(H_t)]$

The constant 1.53 has been so chosen that the standard deviation of G_{ct} is very nearly equal to the standard deviation of A_{ct}. This is needed, for later A_{ct} and G_{ct} measures are used in the same calculation and called M_{ct} measures.

Bearing of the Various Tests Upon English

In the following tables which give the data for the calculation of the regression equations for English with the (1) algebra, English and history tests, and (2) geometry, English and history tests, the probable errors of the calculation involving the algebra and geometry tests are given:

Since the following coefficients of correlation ± their respective probable errors overlap,

r_{EA_t} and r_{EG_t};

$r_{E_tA_t}$ and $r_{E_tG_t}$;

$r_{H_tA_t}$ and $r_{H_tG_t}$;

	E	A_t	E_t	H_t
A_t	.35±.05			
E_t	.44	.37±.05		
H_t	.40	.27±.06	.40	
σ's		3.917±.239	3.274	5.440

and since the standard deviations of A_t and G_t ± their probable errors overlap, A_t and G_t may be combined into a single mathematics group, M_t, without materially affecting the regression equation.

	E	G_t	E_t	H_t
G_t	.31±.07			
E_t	.44	.42±.07		
H_t	.40	.20±.08	.40	
σ's		3.860±.298	3.274	5.440

The regression equation has, accordingly, been calculated from the following table, in which the coefficients of correlation involving the algebra test and geometry test are weighted averages of the coefficients for the algebra and geometry tests separately:

From this table are obtained the following standard deviations:

$\sigma^2_{E.\ M_tE_tH_t} = .722\,\sigma^2_E$

$\sigma^2_{E.\ M_t+E_t+H_t} = .729\,\sigma^2_E$

	E	M_t	E_t	H_t
M_t	.34			
E_t	.44	.38		
H_t	.40	.25	.40	
σ's	5.254	3.896	3.274	5.440

There is so little difference between these standard deviations that the simpler method is used, i.e. $E = \frac{M_t+E_t+H_t}{3}$, designated by E_{ct}.

Bearing of the Various Tests Upon History

By parity of reasoning, the same English combination measure, E_{ct}, is used to correlate with history. The data bearing upon the problem are as follows:

$\sigma^2_{H.\ M_tE_tH_t} = .787\,\sigma^2_H$

$\sigma^2_{H.\ M_t+E_t+H_t} = .793\,\sigma^2_H$

	H	M_t	E_t	H_t
M_t	.37			
E_t	.35	.38		
H_t	.31	.25	.40	
σ's	5.156	3.890	3.136	4.796

From which it may be deduced[1] that $r_{HH_{ct}} = .455$. (This value is used in the calculation Appx. p. 105, but care should be exercised in using values obtained in this way as it should be noticed that errors are cumulative and an error introduced here by throwing away .001's, or for other reasons, may affect a subsequent correlation considerably.)

[1] Same method used as in paragraph upon Grading of History Test, Appx. p. 99.

Interest Tests—Grading of Books

In expressing the grades of the books 1, 2,....7 for English, as deviations from the mean, a normal distribution was assumed for the grades given by each of the judges—this method seemed to be necessary as the means for the different judges varied appreciably—and use was made of the same kind of a transformation as in expressing literal grades in numerical terms.

In the case of the history grading, it does not seem reasonable to assume a normal distribution, as a straight history is surely a greater distance above the mean, as evidence of interest in history, than is a book like "Kite Flying for Boys" below the mean. No simple method is at hand to tell the nature of the distribution of the books with reference to their historical significance, but the assumption of a skewed distribution of some sort is surely more reasonable than the assumption of a normal distribution. A distribution skewed .75 was assumed, skewness being measured by the formula $\frac{3(\text{Av.-Median})}{\sigma}$.[1] With a distribution so skewed, grades 1, 2....7 were expressed as deviations from the mean in the same manner as was done for the grading of books for English, and for literal grades, under the assumption of a normal distribution.

It may be mentioned that the calculation of the reliability coefficient for the grading of books for history was done before the above-mentioned transformation was made, so that a small inaccuracy is present. Calculation of this coefficient after the transformation would raise its value a little above the obtained value, .720.

In grading sports, entertainments, words and magazines, it is not necessary to resort to a transformation, for even though the means should be quite different for different judges, a simple average may be taken, for the standard deviations of the grading of the different judges were found to be very nearly equal, and every judge graded every item except in the case of magazines, where the means, as well as the standard deviations, were nearly equal. This does not introduce the error that would have resulted from such a procedure in the case of books, where the books that were graded by only two judges were not, in many cases, graded by the same two.

Grading of the Interest Tests with Reference to (a) English, (b) Mathematics and (c) History

The data in the following table serve as a basis for combining the various parts of the interest test into a single measure to correlate with English.

	E	Spts.	Ents.	Vocs.	F.of A.	Wds.	Mags.	Bks.
Sports	.20							
Entertainments	.14	.5						
Vocations	.24	.3	.1					
Factor of accuracy	.04	.2	.1	.1				
Words	.26	.3	.0	.1	.7			
Magazines	.37	.4	.3	.4	.2	.2		
Books	.13	.3	.0	.2	.0	.3	.3	
σ's		1.965	3.512	2.059	4.286	4.161	2.232	2.428

[1] The actual distribution used is that represented in Thorndike, Mental and Social Measurements, Ed. 2, p. 74, but any reasonable distribution with a skewness of .75 would yield comparable results.

The coefficients of correlation in this table which involve English were calculated by the Pearson formula, and the balance by the percentage of unlike signs method.

The standard deviations given in this table are not the standard deviations of the original measures. Certain of the measures were grouped for convenience in handling: The following relations hold between the above standard deviations and the standard deviations of the original measures:

1.965 =	Standard	deviation	of original	grading of	sports.
3.512 =	"	"	"	"	entertainments.
2.059 =	"	"	"	"	vocations.
4.286 = 20	"	"	"	"	factor of accuracy.
4.161 = .2	"	"	"	"	words.
2.232 = 3.3	"	"	"	"	magazines.
2.428 = 3.3	"	"	"	"	books.

The labor involved makes it plainly out of the question to calculate a regular regression equation, so that a rough approximation only has been attempted. Since the grading of the magazines correlates the most highly with English, the effect of that one factor is taken into account by calculating partial coefficients of correlation of the type $r_{\text{E Spts} \cdot \text{Mags}}$, $r_{\text{E Ents} \cdot \text{Mags}}$, etc. All such partial coefficients of correlation are given in the accompanying table except $r_{\text{E F of A} \cdot \text{Mags}}$. The factor of accuracy is very highly correlated with the grading of the words, and is therefore evaluated in connection with it.

$r_{\text{E F of A} \cdot \text{Wds}} = -.206$
$r_{\text{E Wds} \cdot \text{F of A}} = .325$

Probably the average of $r_{\text{E Wds} \cdot \text{Mags}}$ and $r_{\text{E Wds} \cdot \text{F of A}}$ gives a better weight than either alone.

$r_{\text{E Spts} \cdot \text{Mags}} = .061$
$r_{\text{E Ents} \cdot \text{Mags}} = .033$
$r_{\text{E Vocs} \cdot \text{Mags}} = .108$
$r_{\text{E Wds} \cdot \text{Mags}} = .204$
$r_{\text{E Bks} \cdot \text{Mags}} = .021$

This gives .265 as the weighting for Wds. and proportionately −.168 for F of A.

The weight assigned to magazines is the average of the following partial coefficients of correlation:

$r_{\text{E Mags} \cdot \text{Spts}} = .331$
$r_{\text{E Mags} \cdot \text{Ents}} = .348$
$r_{\text{E Mags} \cdot \text{Vocs}} = .308$
$r_{\text{E Mags} \cdot \text{Wds}} = .336$
$r_{\text{E Mags} \cdot \text{Bks}} = .350$

average = .338

Having weighted Mags .338, proportionate weightings for other variables are as follows:

$\frac{.338}{.331} = \frac{X}{.061}$, X = weighting for Spts. = .062
Similarly " " " Ents. = .032
" " " Vocs. = .119
" " " Bks. = .020

Since Spts. and Ents. are highly correlated, the weights assigned to them above are somewhat too high. A small amount is arbitrarily deducted from the above weights. The final weights assigned are as follows:

	W'ts
Spts	.058
Ents	.026
Vocs	.119
F of A	−.168
Wds	.265
Mags	.338
Bks	.020

These weights, divided by the respective standard deviations, give the multipliers of the various measures used to obtain a single interest and information test grade. The single grade for the interest and information test in its bearing on English (E_i) is accordingly given by the following equation: (In the following equation the letter E indicates that the grade assigned is in relation to English.)

$$.1515\,E_i = .0295(E_{Spts}) + .0074(E_{Ents}) + .0578(E_{Vocs}) - .0392(\text{F of A}) + .0637(E_{Wds}) + .1515(E_{Mags}) + .0082(E_{Bks})$$

(.1515 is taken for convenience, to make the coefficient of E_{Mags} equal to unity.)

$$\therefore\ E_i = .195(E_{Spts}) + .048(E_{Ents}) + .382(E_{Vocs}).259 - (\text{F of A}) + .420(E_{Wds}) + 1.00(E_{Mags}) + .054(E_{Bks}),\ \text{or, for practical purposes,}$$

$$E_i = .2(E_{Spts}) + .0\tfrac{1}{2}(E_{Ents}) + .4(E_{Vocs}) - .2\tfrac{1}{2}(\text{F of A}) + .4(E_{Wds}) + 1.0(E_{Mags}) + .0\tfrac{1}{2}(E_{Bks})$$

The same relative weighting is assumed in obtaining a single measure of the interest and information test to correlate with history (H) and with mathematics (M), except that the F of A is weighted differently, since $r_{H\,\text{F of A}} = .094$ and $r_{M\,\text{F of A}} = .173$. The following weighting is used.

$$H_i = .2(H_{Spts}) + .0\tfrac{1}{2}(H_{Ents}) + .4(H_{Vocs}) - .1(\text{F of A}) + .4(H_{Wds}) + 1.0(H_{Mags}) + .0\tfrac{1}{2}(H_{Bks})$$

$$M_i = 2[.2(M_{Spts}) + .0\tfrac{1}{2}(M_{Ents}) + .4(M_{Vocs}) + .0\tfrac{1}{2}(\text{F of A}) + .4(M_{Wds})]$$

The factor 2 in this last equation is simply for the purpose of obtaining a more convenient distribution—by maintaining distributions with 20 or more divisions it is not necessary to correct for grouping or for arbitrary means.

Combination of Parts of Interest Test M_i, E_i, H_i, with Reference to (a) Mathematics, (b) English, (c) History

(*a*) *Mathematics.*

Although the interest test has been graded specifically with reference to mathematics, it may be that the gradings with reference to English and history give some light upon the most probable mathematical standing of the pupil.

The calculation of the regression equation involving M, M_i, E_i, and H_i will decide the question. The data for this calculation follow:

From this $\sigma^2_{M.M_iE_iH_i} = .911\sigma^2_M$,
from which it may be deduced that
$r_{M(M_iE_iH_i)} = .298$
This correlation is considerably above the correlation for r_{MM_i}, which equals .24, and by inspection of the table, it is apparent that it is considerably above $r_{M(M_i+E_i+H_i)}$, for the large standard deviation for history (6.12), operates, when taking an average or sum, to weight the history grading the highest, so the regression equation method is plainly the method needed. Calculation gives the following:

	M	M_i	E_i	H_i
M_i	.24			
E_i	.20	.21		
H_i	.15	.54	.63	
σ's	4.93	4.64	3.13	6.12

$$M = .2238\frac{4.701}{3.730}M_i + .1826\frac{4.701}{2.340}E_i - .095\frac{4.701}{3.990}H_i$$
$$= .2822\ M_i + .3669\ E_i - .1120\ H_i$$

Multiplying by the convenient factor, .5646, and designating the result by M_{ci} (combination of the parts of the interest test with reference to mathematics) gives the following:

$M_{ci} = .500\,M_i + .650\,E_i - .199\,H_i$, or for practical purposes, $M_{ci} = .5\,M_i + .65\,E_i - .2\,H_i$.

(*b*) *English.*

Similar data with reference to English are:

From this
$\sigma^2_{E.M_iE_iH_i} = .783\ \sigma^2_E$
from which it may be deduced that
$r_{E(M_iE_iH_i)} = .464$,

	E	M_i	E_i	H_i
M_i	.15			
E_i	.46	.21		
H_i	.32	.54	.63	
σ's	5.23	4.64	3.13	6.12

but as $r_{EE_i} = .46$ there is practically no object in using the longer method. E_i is therefore taken as the sole measure of the interest test in its bearing upon English, and for such use will be designated E_{ci}.

(*c*) *History.*

The data referring to history are as follows:

From this
$\sigma^2_{H.M_iE_iH_i} = .875\ \sigma^2_H$
from which it may be deduced that
$r_{H(M_iE_iH_i)} = .353$

	H	M_i	E_i	H_i
M_i	.02			
E_i	.27	.21		
H_i	.30	.54	.63	
σ's	5.10	4.64	3.13	6.12

It is evident that $r_{H(M_i+E_i+H_i)}$ is appreciably lower than this, for the rather large standard deviation for M_i (4.64) would operate to weight M_i quite heavily. The correlation .353 is sufficiently higher than the correlation r_{HH_i} (.30), to make the regression equation desirable. By calculation,

$$H = -.1597\frac{4.891}{3.778}M_i + .0767\frac{4.891}{2.369}E_i + .2319\frac{4.891}{3.901}H_i$$
$$= -.2070\,M_i + .1585\,E_i + .2910\,H_i$$

To obtain a convenient distribution and a simpler equation to work with, this equation has been multiplied by 2.406. The result is designated by H_{ci}.

$H_{ci} = -.498\,M_i + .381\,E_i + .700\,H_i$, or for all practical purposes, $H_{ci} = -.5\,M_i + .38\,E_i + .7\,H_i$.

COMBINATION OF THE MATHEMATICS TESTS, M_{ct} AND M_{ci}, WITH REFERENCE TO MATHEMATICS. SIMILAR COMBINATIONS OF ENGLISH AND HISTORY TESTS

(*a*) *Mathematics.*

On page 99 of the Appendix, mathematics, English and history tests, M_t, E_t, H_t, were combined into a single grading, M_{ct}, which gives the total bearing of these tests upon mathematics. On page 104 of the Appendix, the gradings of the mathematics, English and history interest tests, M_i, E_i, H_i, are combined into a single grade, M_{ci}, which gives the total bearing of the three interest tests upon mathematics. It now remains to combine M_{ct} and M_{ci} into the single measure which correlates the highest with M. This single measure will be designated by M_c, and is given by the following regression equation, which is based upon the accompanying data:

$$M = .1606\,\frac{4.53}{1.719}\,M_{ci} + .4198\,\frac{4.53}{2.958}\,M_{ct}$$

Multiplying by the convenient factor 1.556 gives:

$M_c = .658\,M_{ci} + 1.00\,M_{ct}$, or, for practical purposes,

$M_c = .66\,M_{ci} + M_{ct}$.

	M	M_{ci}	M_{ct}
M_{ci}	.30		
M_{ct}	.48	.35	
σ's	5.23	1.86	3.48

This relative weighting is used whether it is desired to combine the grades of the tests with reference to algebra or geometry. If the derivation of regression equations for algebra and geometry had been undertaken separately, the difference from the above weighting would have been slight and the increased correlation due to the more exact weighting would have been inappreciable. The terms A_c and G_c will be used instead of M_c, when it is desired to speak of the algebra combination, and the geometry combination, rather than the mathematics combination.

(*b*) *English.*

Data for English, similar to the above for mathematics, are given in the accompanying table:

	E	E_{ci}	E_{ct}
E_{ci}	.46		
E_{ct}	.46	.34	
σ's	5.23	3.21	3.23

Since $r_{EE_{ci}} = r_{EE_{ct}}$, and $\sigma_{E_{ci}}$ is very nearly equal to $\sigma_{E_{ct}}$, a straight average or sum of the two measures is the desired combination, i.e., $E_c = E_{ci} + E_{ct}$.

(*c*) *History.*

The data for history are given in the accompanying table:

Calculation gives:

$$H = .2154\,\frac{4.446}{4.420}\,H_{ci} + .3782\,\frac{4.446}{2.813}\,H_{ct}$$

$$= .2168\,H_{ci} + .5979\,H_{ct}$$

	H	H_{ci}	H_{ct}
H_{ci}	.33		
H_{ct}	.45	.33	
σ's	5.10	4.80	3.23

Multiplying by the convenient factor, 1.673, and designating the result by H_c, gives:

$H_c = .1814\ H_{ci} + 1.00\ H_{ct}$, or, approximately, $H_c = .2\ H_{ci} + H_{ct}$.

Combination of Elementary School Standing, Teachers' Estimates and Test Standing with Reference to Average Class Standing

The estimation of average class standing based upon all three sources of data is by means of the following regression equation, based upon the accompanying data:

	F_A	$7,6,5,4_A$	Est_A	T_A
$7, 6, 5, 4_A$	.83			
Est_A	.81	.68		
T_A	.51	.56	.54	

$$F_A = .6422 \frac{.4458\,\sigma_{F_A}}{.5340\sigma_{(7,\,6,\,5,\,4_A)}} (7,\ 6,\ 5,\ 4_A) + .5983 \frac{.4458\,\sigma_{F_A}}{.5666\sigma_{Est_A}} Est_A$$

$$- .0771 \frac{.4458\sigma_{F_A}}{.7973\sigma_{T_A}} T_A.$$

The equation is left in this form for reference to it. The values of the various standard deviations are:

$\sigma_{F_A} = 3.662$; $\sigma_{(7,\,6,\,5,\,4_A)} = 4.665$; $\sigma_{Est_A} = 6.660$; $\sigma_{T_A} = 3.845$.

SECTION 12

GRADE AND TEST DATA

SECTION 12—GRADE AND TEST DATA

Group 1	Grade in High School Courses											Teachers' Estimates					Tests				
	Test Courses																*Subject Tests*			*Interest Tests*	
	2nd ½ year				1st ½ year																
Individual No.	M.—Geom.	E.	H.—3rd or 4th yr.	H.—2nd yr.	M.—Geom.	Eng.	Hist.	Caesar	Lat. Pr.	Fr.	Ger.	Miscel.	I. a.	Cons.	Emo. i.	Exp.	M_t, $=G_t$	E_t	H_t	M_i, $=G_i$	E_i	H_i	Age
1	A+	B			A+	B		A+	A	B		C+	5	7	4	−4	−2	−4	6	−4	1	1	15.5
2	B	B		C	C	B		C	D	C	C		−6	0	−3	−3	−1	−2	4	−3	−2	−5	14.0
3	D	D			C+	C	B	C	C+				0	−8	0	2	−3	4	4	−7	−5	−12	15.4
4	B+	B			B	B	A++			E		B++	8	12	12	8	2	0	4	3	7	8	16.2
5	D	D	D	C+	C	D	D++		C+			D++	−6	0	−6	−4	−3	−3	−8	2	4	4	17.0
6	C	C	C	D+	B	C	C+				B	A	1	−8	−1	−2	−2	1	−8	−3	−1	11	16.8
7	C+	D			D	C	B						2	2	−5	−1	4	5	4	10	5	7	17.1
8	B	C		B	C	B	D		C+				0	−4	−2	−6	−4	1	−5	−2	4	11	15.0
9	D−	B		D+	B	B	C+		E			D−	6	3	8	6	−4	0	6	−1	−1	−8	16.3
10	C	D		C	D	D	B		C	C+			−5	−1	0	−2	−1	−1	−5	−8	−2	−1	15.1
11	D	D	C		C	C	D		D				−2	8	2	2	−4	−5	6	8	−2	4	19.1
12	B	B	B		A	D	C+		B		B		6	6	4	2	−5	−5	4	2	−3	−4	15.9
13	C	B			D	C			C	D			−2	−7	−8	−2	−4	−1	4	−5	−1	−10	14.5
14	D	C			A	C			E	C			12	10	11	10	−4	−1	8	−8	−7	−8	17.3
15	A+	C+			A	A			C	B			6	12	13	4	4	6	3	8	7	16	15.6
16	A	A			A	A			D	B			−2	−5	−3	−4	4	3	−6	1	−3	6	15.0
17	D	D			A	D			E	C		B	2	2	−4	6	0	1	−5	8	2	−1	15.6
18	D+	B			C	B			B	B		D+	0	2	2	2	4	2	5	4	3	6	15.0
19	B	C			B	C			D	C	D	C	−10	−2	−8	−8	−1	−3	2	−1	2	2	15.4
20	D	B	E−		E	B			E	D	E	E	−1	−2	0	0	−2	2	5	−5	−2	1	17.5
21	D	C	C+		D+	C	D		C+	C+												−6	15.6

GRADE AND TEST DATA—*Continued*

Group 2	Individual No.	Grades in High School Courses												Tests						Age
		Test Courses												Teachers' Estimates				*Subject Tests*		*Interest Tests*	
		2nd ½ year				1st ½ year															
		M.—Alg.	E.	H.—3rd yr.	H.—2nd yr.	M.—Alg.	Eng.	Hist.	Lat.	Fr.	Ger.	Miscel.		I. s.	Cons.	Imo. i.	Exp.	$M_t = A_t$	E_t	$M_i = A_i$	E_i	H_i	
	60	D	C		B	D	C	B	D	C		B		0	0	−1	−4	1	1	6	1	3	14.9
	61	D	C		C	E	+	+	+			C		−5	6	−4	−6	−7	−2	−1	−4	−3	14.6
	62	C	+		A	C	C	+				D		−6	−5	10	−10	3	1	−1	−4	3	14.0
	63	+	B		B	A	C	+	C			+		5	0	0	2	4	2	3	1	−3	13.2
	64	C	+		+	C	C	+	C			B		0	0	0	0	4	−4	2	2	3	15.3
	65	C	B		D	D	C	+	D			+		−4	0	6	−3	4	0	2	−2	−4	14.1
	66	+	D		D	C	C	+	−		D			−2	0	−2	0	−2	1	12	2	4	14.1
	67	C	C		D	C	C	+	+	C				−8	−4	−5	−2	−5	−8	0	−2	−6	16.0
	68	+	B		D	C	D	+	E					−8	−8	−5	−8	−5	4	−1	1	0	14.5
	69	C	C		B	E	D	+	D	C		C		0	−3	0	0	−3	−2	−5	−1	8	14.8
	70	C	+		D	C	C	+	E			+		4	4	−2	6	1	−2	−6	−1	−6	15.7
	71	D	−		D	D	B	+	D	C		D		−2	−3	−2	2	−2	−3	−5	−2	−7	14.4
	72	B	C		D	B	B	+	E			C		6	2	−2	6	3	2	−1	−2	2	14.2
	73	C	+		D	B	C	+	D			B		0	0	0	0		1	−3	4	2	14.1
	74	B	+		D	C	B	+	E					4	11	5	14		1	0	2	0	13.5
	75	E	+		D	D	C	+	−	E	D			−10	−10	−10	−8	−1	−7	0	−1	−2	15.8

GRADE AND TEST DATA—*Continued*

Group 3	Grades in High School Courses										Teachers' Estimates				Tests						Age
	Test Courses														*Subject Tests*			*Interest Tests*			
	2nd ½ year			1st ½ year																	
Individual No.	M.—Alg.	E.	H.—2nd yr.	M.—Alg.	Eng.	Sci.	Lat.	Fr.	Ger.	Miscel.	I. a.	Cons.	Emo. i.	Exp.	M_t, $=A_t$	E_t	H_t	M_i, $=A_i$	E_i	H_i	
102	D+	C		C	D+	C	D		D		0	−2	−2	−2	−3			6	−2	5	13.7
103	E+	B		D+	B	C	D−			C	−2	−10	−2	0	−7	2	2	−4	6	4	13.4
104	E+	D+		E+	C	D	D+	D+			0	−6	4	0	0	−2	−4	1	−2	2	14.9
105	A	B		B	B+	B+	C	B+		B+	0	0	0	1	5	4	−10	1	5	−1	12.4
106	D−	D+		C	D+	C+	C			B	0	0	0	−2	−2	1	−4	6	−3	2	14.4
107	B	D+		B	C+	C	C+				4	0	−2	0	3	1	−7	0	0	−4	12.9
108	B+	A		A	A	A	A		A	C	12	12	12	10	3	6	7	6	5	10	12.3
109	B	C		C	C+	C	D+			C+	0	0	0	−1	6	1	−4	4	−1	−3	14.2
110	A	A		A	A	B+	A		A		12	11	12	10	4	7	3	3	4	10	12.8
111	D+	C+		D+	C+		D−		D+		2	−6	−9	−2	3	−3	3	−2	−1	−5	13.5
112	D	C		D+	C	C		C	C	D+	−6	0	0	−11	−5	0	3	−4	4	0	13.0
113	C+	C		B+	C	C	C				0	4	0	−4	2	−1	−3	2	−2	−6	14.4
114	C+	A		B+	B+	C+	B	D		C+	4	4	4	4	0	4	4	0	4	2	14.6
115	C	D+		C+	C		D−		D	B	0	−2	0	−6	5	1	−1	3	−3	−5	13.0
116	B	C		B	C				B	C+	−5	4	−8	−4	0	0	1	−4	−5	−6	14.5
117	E+	D+		D+	C		D−			D	0	−4	−2	−4	6	−4	1	2	−1	−2	11.9
118	E	D		B	C	D	E			B	−2	−4	2	−7	−7	1	0	−3	−3	−4	13.8
119	D+	D+		C	C+	C	C			C+	−2	3	0	−2	−3	0	−2	−2	−1	−5	14.0
120	A	A		A	A	B	A			B	11	7	0	10	−1	5	−2	6	5	15	13.4
121	E−	D		D	D+	C+	E			C+	−8	6	0	−8	−5	−7	−2	4	2	5	16.6
122	C	C+		C+	B	C	C	D+		C+	0	−4	−8	−1	−1	−2	−1	−3	−2	−3	14.7
123	C	D+		B+	C	C	D				−6	0	4	−6	0	−2	−8	0	−5	−10	14.0
124	D	D		E+	D+					D	−12	−12	−8	−11	−4	−7	−9				14.7
125	E+	D+		D	C	B	D+			C+	0	0	0	−3	−2	5	5	1	1	6	13.6
126	D−	C		D+	C+	C+	D		B	B	2	2	2	2	−10	−8	0	3	−2	3	13.4
127	C	C		C	D+	D	D+			C	−11	−6	−6	−7	−3	−1	−1	0	−8	−8	13.2
128	B	D+		B+	C+	C+		D+			4	6	6	2	0	2	−5	3	−3	−2	14.1
129	A	A		B+	A	B	B+	A		C+	12	8	0	8		3	12	−5	2	−6	14.5
130	C	D+		D	C+	C	D	D			−6	0	0	0	0	−1	6	−1	−1	0	14.7
131	B	C		C+	C+	B	D+	D			0	0	0	4	−7	−7	2	5	−1	5	15.1
132	D+	C+		D+	C+	D+	D+		D		0	−2	−2	−5	1	0	3	4	−1	9	13.5

GRADE AND TEST DATA—*Continued*

Group 4	Grades in High School Courses										Teachers' Estimates				Tests				
	Test Courses														*Subject Tests*			*Interest Test*	
	2nd ½ year		1st ½ year																
Individual No.	M.—Geom.	E.	M.—Geom.	Eng.	Lat.	Fr.	Ger.	Alg.	Miscel.		I. a.	Cons.	Emo. i.	Exp.	$M_{i'}=G_{i'}$	$E_{i'}$	$H_{i'}$	$M_{t'}=G_{t'}$	$E_{i'}$	$H_{i'}$	Age
183	B	B+	C+	B	C	A	C½−	B+	C						3	−4	7	1	0	1	15.2
184	C	D+	D	B	C		C½−	C	C						−7	−1	4	1	−1	6	15.4
185	A	C	C	B	B+	A	B+	D+	C						4	1	4	6	0	3	14.0
186	C	A	A	B	C			C	B						3	2	1	3	3	2	14.8
187	C	B	C	B	C			A	B						−4	2	1	1	5	10	15.3
188	C	C	C	C	C			C	B						6	3	−3	0	−2	10	15.7
189	B	B	C½−	B	C	D+		D+	D++						−4	−3	−1	−2	−2	5	17.8
190			B	B+	B+	B+		C½−	D++						−1	1	−10	−3	0	−14	14.3
191	B	A	B					C	C						2	1	−5	0	5	0	15.9
192	B	B	B	B				B	B	A					5	1		3	0	−4	17.7
193	C	B	C	B	C	A		D	D	C					1	1	11	−6	1	−2	16.2
194	B	A	C	B	C	B	C½−	C	B	B					−1	3	2	−3	2	1	15.6
195	C	B	B	B	B+		C	B	B	B					3	2	5	−2	6	−7	
196	D	C	D	B	C	A+	D	C½−	C½−	C					2	0	4	−5	2	−1	16.7
197	C	B	C	B	C		B	B	B	B					3	−1	3	2	−1	0	14.6
198	D	C	D	B	C	A	C	C	B	B					−5			−5	1	−1	16.4
199	B	C	C	B	C	B		B	C	C					−3	0	−2	1	−1	−1	15.3
200	C	B	C	B		A		C	A	B					−2	5	−4	11	−4	−1	16.3
201	C	B	C	B	C	B		D	C						−5	1	−1	3	8	6	16.3
202	D	C	D	B	C		D++	C+							1			0	0	10	15.6
203	B	B	C½−	B	C	A	D	B	B	B					1	−6	−11	−2	−1	−1	15.3
204	C	B	B	B	C			B	B	B					−3	1	0	−3	−2	−4	
205	D	B	D	B	C	A	C++	D	B	B					3	−2	2	−4	−4	1	15.8
206	D	D+	D	B+	C		B++	D	B	B					−4	−3	1	0	−2	−6	15.7
207	C	B+	D	C½−	C½−		C	B	C	B					−2	−2	5	−3	−3	1	15.2
208	D+	D+	C½−	C½−	C½−		C	B	B						4	−1	−4	−5	−3	−9	15.2

GRADE AND TEST DATA—*Concluded*

Group 5	Grade in High School Courses													Teachers' Estimates				Tests						Age
	Test Courses																	Subject Tests			Interest Tests			
	2nd ½ year		1st ½ year																					
Individual No.	M.—Alg.	E.	M.—Alg.	Lit.	Comp.	Gram.	Hist.	Ger.	Geog. and N. H.	Art. and Man. T.	Music	Phys. Tr.	Dom. Art.	I. a.	Cons.	Emo. i.	Exp.	$M_t, = A_t$	E_t	H_t	$M_i, = A_i$	E_i	H_i	
209	C½–	C	C	C	C	C	C½–	C½–	B	B	C+	C+						1	1	8	3	–1	2	13.5
210	B	C	C	C	C	C½–	C		B	A½–	C	B						–1	3	5	7	–3	2	14.5
211	B	B	C	B	C+	B	B	C+	C+	C+	B	B							3	9	1	3	–2	13.6
212	C+	B	C	C+	B	B+	C+		C	B	C	C							2	11	–3	2	3	16.6
213	C	B	D+	C	C+	B	C+	C+	C+	D+	B	C+	C					2	0	–3	–7	–1	4	14.4
214	C	C	C	C	C½–	C	D+	C	C+	B	C½–	B						–4	–5	–3	3	–1	6	13.9
215	C+	D+	D	C+	D+	B½–	D+	D+	C+	C	C½–	C+						–4	–1	–3	–8	–4	–7	14.7
216	B	A		B	B+	A	C+	B	B+	B	A	B	B					3			5	2	3	13.0
217	C	C	D	C+	C+	C½–	C½–		B	C	B	B	C					–1	4	–11	–2	1	5	14.0
218	B	C	B	C+	C	C	C	C+	C+	C	C	C						6			7	1	1	12.6
219	C½–	B	D	B	B	C	B+	C+	C+	C+	C	C						1	5	9	0	4	6	12.3
220	B	C	B	C	C	C	C½–	C+	B	A	A	C						0	–2	–5	–3	–4	–8	13.8
221	B	B	B	C+	C+	B+	C+	B	C+	B	C+	B						–2	–2	–1	–1	–1	2	14.0
222	C	C	D	B	B	C½–	B	B	B	C	B	B	B					–2	–3	–14	1	1	1	13.9
223	C	C	C	C+	C+	C+	C½–		B	A	C	B						3	–2	–1	–1	0	–2	16.8
224	C+	B	D+	C+	C	B	C	B	B+	C+	B	C+	C					2	2	2	–2	–1	–2	13.7
225	C	B	C	B	C+	C½–	C+	C½–	C+	A	C+	B						2	2	4	–3	1	4	13.4
226	C+	C	D+	C+	C+	C+	C		C+	C	B+	B	B					–4	–6	–10	2	0	–2	15.3
227	C½–	D+	D	C½–	C½–	D	D	D	B	D+	B	C	C					0			–1	–5	4	17.2
228	B+	B	B	B+	C+	C+	C	C+	B+	B	B	C+	B					5	1	9	4	–1	2	14.3
229	C	D+	B	C	C	C	C+	C	B+	C½–	B	C+	C					1			–3	–1	–7	15.2
230	B	B	B	C+	C+	C	C		B+	C+	B	C+	B					0	–6	–11	–1	0	0	16.3
231	B	B	C	C+	C+	C+	B	C	B+	C+	B	A	B					–6	3	9	–1	0	–2	13.3
232	C+	B	C	B	B	C	C	C	B	C	B	B	B					–1	2	3	–3	0	–5	15.0
233	A	A	C	B+	B+	C+	C	C	A	C	B	A	A					1	0	1	5	1	1	15.9

HIGH AND ELEMENTARY SCHOOL GRADES

EXPRESSED AS DEVIATIONS FROM MEAN

Individual No.	High School 1st year			7th Gr.			6th Gr.			5th Gr.			4th Gr.		
	F_M	F_E	F_{AV}	7_M	7_E	7_H	6_M	6_E	6_H	5_M	5_E	5_H	4_M	4_E	4_H
4	3	−1	0	0	3	−1	1	1	0	1	2	1	0	−1	1
9	−10	−10	−5	1	2	0	1	3	1	1	3	1	0	1	1
10	1	−1	1	−2	−3	0	−1	0	0	1	−1	1	0	4	−0
12	3	−4	3	0	0	−1	−4	0	−1	−3	−3	−2	1	2	1
14	−4	−6	−5	−2	3	2	0	7	1	0	6	4	−2	8	−3
15	10	8	9	−3	−4	−4	−2	−6	−3	0	−3	0	−2	−4	−2
16	−1	4	3	1	−3	−2	2	−3	0	1	−2	0	0	−3	−1
17	1	−2	−2	0	1	1	2	3	2	1	3	1	1	3	3
18	−3	−1	−2	1	2	1	3	5	2	0	0	0	0	2	−2
24	10	8	7	0	−1	−3	−4	−6	−2	0	−5	−2	−4	−6	−3
25	7	4	5	0	3	−3	−1	3	2	−3	3	−2	−4	−3	−3
29	7	4	5	−3	−2	−2	−3	−4	−2	−1	2	0	−3	−4	−1
36	3	4	0	−3	−3	−2	−2	3	1	0	3	2	−4	2	−2
38	3	8	4	−2	−5	−1	−3	−5	0	2	−3	1	−2	−4	−2
39	5	0	−1	0	1	−2	0	3	−2	1	2	1	1	2	0
40	−6	4	0	0	0	1	1	−1	0	−4	−5	−1	−4	−5	−1
41	0	−1	−3	2	0	−1	1	−2	0	0	3	2	2	0	4
42	−1	−10	−6	2	4	0	−1	3	1	−3	2	1	−1	3	1
45	7	1	1	2	−1	−1	−3	0	−2	−3	−3	−2	−4	−6	−3
46	−4	−2	−3	1	−1	1	−2	−1	0	1	0	1	−2	1	−3
47	5	1	2	0	2	0	−1	−1	0	0	3	0	−1	2	1
53	5	4	6	−3	−3	−3	−3	−3	−2	−1	0	−2	−2	−3	−3
65		−2	0	2	0	1	2	2	−1	−1	0	1	3	2	1
76		−6	−5	1	1	0	1	1	1	−3	−2	−1	−1	3	1
89		−2	−3	1	2	0	2	−3	1	3	−3	1	1	0	1
98		−1	−2	2	3	2	4	2	2	1	3	3	5	5	4
103	−4	3	−1	2	−1	0	1	−2	−3	0	−2	−2	2	−3	−1
106	−3	−6	−3	0	0	0	0	1	0	0	−2	1	2	−2	1
109	2	0	0	0	3	−1	−1	−1	0	1	−2	1	−1	−3	1
110	12	9	10	−1	−2	−3	−3	−3	−2	1	−6	−3	−1	−5	−3
112	−3	−2	−2	1	−2	0	1	2	0	1	0	1	−3	−6	−1
113	4	−2	1	0	6	0	0	5	0	0	6	2	2	2	1
116	4	−2	2	0	−2	−1	1	−1	0	−3	0	−2	−1	−4	−1
123	3	−4	−1	0	6	1	0	4	1	0	3	2	−3	−3	1
125	−5	−4	−4	1	3	0	0	2	−1	−3	1	−2	3	2	1
127	0	−4	−2	−1	2	0	0	3	−2	−3	0	−1	−2	0	1
128	5	−2	0	0	2	−1	−1	0	−1	−1	0	0	−2	0	1
129	8	9	9	−1	−3	−4	0	−5	−2	−3	−6	−3	−1	−4	−1
130	−2	−2	−3	0	2	0	1	2	0	2	1	2	−1	0	−2
135	−1	5	2	−1	3	−1	1	2	0	0	−2	1	−1	−1	−1
138	8	6	6	−2	−1	−1	−4	0	−4	−2	−4	−2	−3	−1	−1
143	6	7	7	−1	−2	−2	−1	−3	−3	−3	−5	−3	−3	−5	−3
145	2	5	3	−1	−1	0	0	0	0	1	−5	0	−3	−5	−1
147	6	0	3	0	3	−1	0	−4	−1	−3	−1	2	−2	−4	−3
148	2	5	3	−2	−2	−1	0	−1	0	−2	−5	−2	−1	−3	−1
151	5	−2	2	0	−2	0	0	0	−1	−1	6	−1	0	2	−1
152	0	1	1	1	2	0	0	0	0	−1	−4	−1	−1	0	−1
153	−1	2	1	−1	2	0	0	1	−1	−3	2	−3	−3	−1	1
154	0	5	3	2	−2	0	3	0	1	1	−6	−1	1	−1	1
156	−2	2	−1	−1	1	0	1	1	0	−1	−6	−1	0	−6	−1
157	4	−2	0	4	−2	1	1	2	−2	0	2	−3	1	−2	1
160	0	1	0	1	−1	0	−1	−3	−1	−3	−4	1	−1	1	1
162	−1	0	−2	0	0	0	0	2	2	0	−2	3	2	−1	2
166	6	0	3	1	1	0	0	2	0	0	−1	1	1	−2	1
168	6	3	4	0	0	1	0	−3	−1	−1	−2	1	−3	−3	1
169	−8	3	−1	1	0	2	1	−3	0	2	−6	−2	2	−6	1
172	−2	−8	−5	0	3	1	0	2	1	−2	−2	2	0	−1	1
173	2	2	1	0	1	−1	−1	1	−1	1	0	0	1	−3	0
174	−5	7	5	0	−4	−4	0	−4	−4	0	−6	−3	1	−4	1

SECTION 12—GRADE AND TEST DATA

Group 1	Individual No.	Grade in High School Courses																Tests				
		Test Courses																Teachers' Estimates			*Subject Tests*			*Interest Tests*		Age
		2nd ½ year				1st ½ year																				
		M.—Geom.	E.	H.—3rd or 4th yr.	H.—2nd yr.	M.—Geom.	Eng.	Hist.	Caesar	Lat. Pr.	Fr.	Ger.	Miscel.	I. a.	Cons.	Emo. i.	Exp.	$M_t = G_t$	E_t	H_t	$M_i = G_i$	E_i	H_i	
	1	A+	B		C	A+	B	B	A+	A	B		C+	5	7	4	−4	−2	−4	6	−4	1	1	15.5
	2	B	B			C	B		C	D	C			−6	0	−3	−3	−1	−2	4	−3	−2	−5	14.0
	3	D	D			B+	C		D	C+	E			0	−8	0	2	−3	4	4	−7	−5	−12	15.4
	4	B	B			C	C	A+	C			B		8	12	12	8	2	0	4	3	4	8	16.2
	5	D	D	C+		B+	B							−6	0	−6	−4	−3	−3	−8	2	−1	4	17.0
	6	C	C	D+		C	C		D	C+	C+			1	−2	−1	−2	−3	−2	8	3	−1	11	16.8
	7	C	B			B	C							2	−8	−5	1	2	5	4	2	5	7	17.1
	8	B	D			D	C							0	2	−6	−6	0	1	4	1	4	11	15.0
	9	D	B			B+	B							6	4	−2	6	−4	6	4	−1	−1	−8	16.3
	10	C	D			D	C							−5	−1	8	0	5	−1	−5	8	0	−1	15.1
	11	D	D			C	C						D−	−6	8	0	−2	−4	−2	6	−4	2	8	19.1
	12	B	B			D	C							6	−8	2	2	5	−1	−8	2	−1	4	15.9
	13	C	B	C		A	A							−2	6	4	−2	−5	−3	8	1	−3	−10	14.5
	14	D	C			C	A							−4	−7	−8	10	−4	−1	−5	−7	−7	−8	17.3
	15	A+	A+			A+	A+							12	10	11	4	8	6	8	8	7	16	15.6
	16	A	A			A	A							6	12	13	4	−4	3	2	1	7	6	15.0
	17	D	D			B	D							−2	−5	−3	−4	4	3	−6	5	−2	−1	15.6
	18	D	D			B	B						B+	2	2	−4	6	0	1	−5	4	3	6	15.0
	19	B	B			C	C						D+	0	2	2	2	4	2	5	−1	2	2	15.4
	20	D	D			E	E					E	C	−10	−2	−8	−8	−1	−2	2	−3	−2	1	17.5
	21	D	C	C+		D	B	D			C+		E	−1	−2	0	0	−2	−4	5	−5	−2	−6	15.6

GRADE AND TEST DATA—*Continued*

Group 2	Individual No.	Grades in High School Courses										Teachers' Estimates				Tests				Age
		Test Courses														*Subject Tests*			*Interest Tests*	
		2nd ½ year			1st ½ year															
		M.—Alg.	E.	H.—3rd yr.	H.—2nd yr.	M.—Alg.	Eng.	Hist.	Lat.	Fr.	Ger.	Miscel.	I. s.	Cons.	Emo. i.	Exp.	$M_t - A_t$	E_t	H_t	$M_i - A_i$	E_i	H_i	
	60	D	C		B	D	C	B	D	C		B	0	0	−1	−4	1	1	0	6	1	3	14.9
	61	D	C		C	E	D	+	+			C	−5	6	−4	−6	−7	−2	−5	−1	−4	−3	14.6
	62	C	C		A	C	C	+				D	−6	−5	10	−10	3	1	−4	−4	1	−3	14.0
	63	B	B		B	B	D	+	+			+	5	0	0	2	4	2	2	3	3	3	13.2
	64	D	D		D	C	C	A	D			D	0	0	0	0	4	−4	−2	2	1	−3	15.3
	65	C	C		C	B	C	C	C			D	−4	0	0	−3	4	0	−2	12	2	4	14.1
	66	B	C		C	B	C	B	D	C		D	0	−4	6	0	−2	1	2	0	0	−6	14.1
	67	C	E		C	E	C	B	D			+	−2	−8	−2	−2	−5	−8	2	−1	−2	0	16.0
	68	C	C		C	C	D	B	D	C		C	−8	0	−5	−8	−3	−5	0	−1	1	−6	14.5
	69	D	D		B	D	B	B	D			+	0	4	0	0	1	4	2	−1	7	8	14.8
	70	D	C		D	D	C	B	D	C		D	4	−3	4	6	−3	−2	−2	−5	−1	−6	15.7
	71	D	D		D	C	D	A	D	C		C	−2	−2	−2	2	1	−2	−3	−5	−2	−7	14.4
	72	B	C		B	B	C	C	D	C		B	6	2	−2	6	3	1	2	1	−2	2	14.2
	73	B	C		B	C	C	C	D			C	0	0	0	0		2	6	−1	4	2	14.1
	74	B	B		C	C	B	B	E			B	4	11	5	14		1	4	−3	2	0	13.5
	75	D	E		D	D	E	D	E			B	−10	−10	−10	−8	−1	−7	0	0	−1	−2	15.8

GRADE AND TEST DATA—*Continued*

Group 3	Grades in High School Courses										Teachers' Estimates				Tests						Age
	Test Courses														*Subject Tests*			*Interest Tests*			
	2nd ½ year			1st ½ year																	
Individual No.	M.—Alg.	E.	H.—2nd yr.	M.—Alg.	Eng.	Sci.	Lat.	Fr.	Ger.	Miscel.	I. a.	Cons.	Emo. i.	Exp.	M_t, $=A_t$	E_t	H_t	M_i, $=A_i$	E_i	H_i	
102	D+	C		C	D+	C	D		D		0	−2	−2	−2	−3			6	−2	5	13.7
103	E+	B		D+	B	C	D−			C	−2	−10	−2	0	−7	2	2	−4	6	4	13.4
104	E+	D+		E+	C	D	D+	D+			0	−6	4	0	0	−2	−4	1	−2	2	14.9
105	A	B		B	B+	B+	C	B+		B+	0	0	0	1	5	4	−10	1	5	−1	12.4
106	D−	D+		C	D+	C+	C			B	0	0	0	−2	−2	1	−4	6	−3	2	14.4
107	B	D+		B	C+	C	C+				4	0	−2	0	3	1	−7	0	0	−4	12.9
108	B+	A		A	A	A	A		A	C	12	12	12	10	3	6	7	6	5	10	12.3
109	B	C		C	C+	C	D+			C+	0	0	0	−1	6	1	−4	4	−1	−3	14.2
110	A	A		A	A	B+	A		A		12	11	12	10	4	7	3	3	4	10	12.8
111	D+	C+		D+	C+		D−		D+		2	−6	−9	−2	3	−3	3	−2	−1	−5	13.5
112	D	C		D+	C	C		C	C	D+	−6	0	0	−11	−5	0	3	−4	4	0	13.0
113	C+	C		B+	C	C+	C				0	4	0	−4	2	−1	−3	2	−2	−6	14.4
114	C+	A		B+	B+	C+	B	D		C+	4	4	4	4	0	4	4	0	4	2	14.6
115	C	D+		C+	C		D−		D	B	0	−2	0	−6	5	1	−1	3	−3	−5	13.0
116	B	C		B	C				B	C+	−5	4	−8	−4	0	0	1	−4	−5	−6	14.5
117	E+	D+		D+	C		D−			D	0	−4	−2	−4	6	−4	1	2	−1	−2	11.9
118	E	D		B	C	D	E			B	−2	−4	2	−7	−7	1	0	−3	−3	−4	13.8
119	D+	D+		C	C+	C	C			C+	−2	3	0	−2	−3	0	−2	−2	−1	−5	14.0
120	A	A		A	A	B	A			B	11	7	0	10	−1	5	−2	6	5	15	13.4
121	E−	D		D	D+	C+	E			C+	−8	6	0	−8	−5	−7	−2	4	2	5	16.6
122	C	C+		C+	B	C	C	D+		C+	0	−4	−8	−1	−1	−2	−1	−3	−2	−3	14.7
123	C	D+		B+	C	C	D				−6	0	4	−6	0	−2	−8	0	−5	−10	14.0
124	D	D		E+	D+					D	−12	−12	−8	−11	−4	−7	−9				14.7
125	E+	D+		D	C	B	D+			C+	0	0	0	−3	−2	5	5	1	1	6	13.6
126	D−	C		D+	C+	C+	D		B	B	2	2	2	2	−10	−8	0	3	−2	3	13.4
127	C	C		C	D+	D	D+			C	−11	−6	−6	−7	−3	−1	−1	0	−8	−8	13.2
128	B	D+		B+	C+	C+		D+			4	6	6	2	0	2	−5	3	−3	−2	14.1
129	A	A		B+	A	B	B+	A		C+	12	8	0	8		3	12	−5	2	−6	14.5
130	C	D+		D	C+	C	D	D			−6	0	0	0	0	−1	6	−1	−1	0	14.7
131	B	C		C+	C+	B	D+	D			0	0	0	4	−7	−7	2	5	−1	5	15.1
132	D+	C+		D+	C+	D+	D+		D		0	−2	−2	−5	1	0	3	4	−1	9	13.5

GRADE AND TEST DATA—*Continued*

Group 4	Grades in High School Courses											Teachers' Estimates				Tests			
	Test Courses															*Subject Tests*			*Interest Test*	
	2nd ½ year		1st ½ year																	
Individual No.	M.—Geom.	E.	M.—Geom.	Eng.	Lat.	Fr.	Ger.	Alg.	Miscel.			I. a.	Cons.	Emo. i.	Exp.	$M_{i'}=G_{i'}$	E_t	H_t	$M_{i'}=G_t$	E_i	H_i	Age
183	B	B+	C+	B	C	A	C½−	B+	C							3	−4	7	1	0	1	15.2
184	D	D C	D C	B C	C		D C	D C	C							−7	1	4	6	−1	6	15.4
185	C	C A	A C	C B	B+	A	B+	C	C B							4	3		3	0	3	14.0
186	A	A B	C A	B C	C		C	A	D+							3	2		6	2	10	14.8
187	C	B C	C C	B C	C			C	D+							−4	2	1	1	5	10	15.3
188	C	C B	C C	B C	C			C	C							6	3	−3	0	−2	5	15.7
189	B	B	C½−	B		D		D								−4	−3	−1	−2	0	−14	17.8
190	B	A	B	B	C½−	B+		C½−								−1	1	−10	−3	5	0	14.3
191	B	B	B	B	B½+		C	C	A							−2	1	−5	0	0	−4	15.9
192	B	B C	C	B	C	A		B	C							5	1	11	3	1	−2	17.7
193	B	B A	C	B	C	B	C½−	B	B							1	−1	2	−6	2	1	16.2
194	C	B C	B	B	B			B	B							−5	3	5	−3	6	−7	15.6
195	C	B C	C	B	B	C+		C½−	B							2	2	4	−2	2	−1	16.7
196	D	C B	D	C	C	A	D+		C							3	0	3	−5	−1	0	14.6
197	C	D C	C	B	B	A		B	B							−5	−1	−1	2	1	−1	16.4
198	D	D C	D	C	B	B		C	B							−3	−3	0	1	−1	−1	15.3
199	B	B C	C	B		B		A	B							−2	2	5	11	−4	−1	16.3
200	C	C B	C	B	C	A		D+	C							−5	−5	1	3	8	6	16.3
201	D	D C	D	C	C	B	D+		C							1	1	−6	0	0	10	15.6
202	B	B C	C½+	C½−	C+		D+	B	B							−3	−3	−11	−2	−1	−1	15.3
203	C	B C	C	B	C	A	D	B	B							−3	−1	0	−3	−2	−4	
204	D	B C	D	C	B		C	B	B							−4	−2	2	−4	−4	1	15.8
205	D	B C	C	B	C	A	C+	B	B							−2	−3	1	0	−2	−6	15.7
206	C	D+	D	C	B		B+	B	C							−2	−3	5	−3	−3	1	15.2
207	C	B+	C½−	C½−	C½−		B+	B	B							4	−1	−4	−5	−3	−9	15.2
208	D	D+	C½−	C½−			C	B	B													

GRADE AND TEST DATA—*Concluded*

Group 5	Grade in High School Courses																			Teachers' Estimates				Tests				
	Test Courses																											
	2nd ½ year								1st ½ year															*Subject Tests*				*Interest Tests*
Individual No.	M.—Alg.	E.	M.—Alg.	Lit.	Comp.	Gram.	Hist.	Ger.	Geog. and N. H.	Art. and Man.	Music	Phys. Tr.	Dom. Art.	I. a.	Cons.	Emo. i.	Exp.	$M_t - A_t$	E_t	H_t	$M_i - A_i$	E_i	H_i	Age
209	C–	C	C	C	C	C	C–	C–	B	B–	C	+						–1	1	8	3	–1	2	13.5
210	B	C	C	C	C	B	C	C	B	A+	C	C	C					–1	3	5	7	–3	2	14.5
211	B	B	C	C	+	+	+	+	C	C+	B	B	B						3	9	1	–3	–2	16.6
212	C	B	C	C	+	B	C	C	C	B	C	C	C					2	2	11	–3	2	3	16.4
213	C	B	C	C	+	C	C	C	C	D	B	C	C						0	–3	–7	–1	4	14.4
214	C	C	D	C	+	B–	C	D	C	+	C–	B	C					–4	–5	–3	–1	–1	6	13.9
215	C	D	+	C	+	A	C	D	B	+	C–	C	C					–4	–1	–3	–8	–4	–7	14.7
216	D	A	C	C	+	C	C	B	C	B	A	B	B					3			5	2	3	13.0
217	A	C	C	B	+	C	C	C	B	C	B	B	C					–1	4	–11	–2	1	5	14.0
218	C	C	B	B	+	C	C	C	B	C	C	B	C					6			7	1	1	12.6
219	B	B	C–	C	+	C	C	C	B	C	C	B	C					1	5	9	0	4	6	12.3
220	B	B	B	C	+	C	C	C	B	C	A	B	C					0	–2	–5	–3	–4	–8	13.8
221	B	C	C	C	+	C	C	C	B	C	B	B	C					–2	–2	–1	–1	–1	–2	14.0
222	C	C	C	C	+	C	C	B	B	C	C	B	C					–2	–3	–14	–1	–1	1	13.9
223	C	C	C	C	+	C	C	B–	B	C	B	B	B					3	2	–1	–1	0	–2	16.8
224	C	C	C	C	+	C	C	C	B	C	C	B	C					2	2	2	–2	–1	–2	13.7
225	C	B	C	C	+	C	C	C	B	C	B	B	B					2	2	4	–3	–1	4	13.4
226	C	B	C	C	+	C	C	C	B	C	B	B	B					–4	–6	–10	2	0	–2	15.3
227	C–	D	D	C	+	C	C	D	C	D	B	C	C					0	1	9	–1	–5	4	17.2
228	B	B	D	B	+	C	C	C	C	+	B	C	C					5			4	–1	2	14.3
229	C	D	B	C	+	C	C	C	C	+	B	C	C					1	–6	–11	–3	–1	–7	15.2
230	B	B	B	B	+	C	C	C	B	+	B	C	C					0	3	9	–1	0	0	16.3
231	B	B	B	C	+	B	C	C	B	+	B	A	B					–6	–1	3	–1	0	–2	13.3
232	C	A	B	C	+	B	C	C	A	+	B	B	B					–1	2	3	–3	0	–5	15.0
233	A	A	C	B	+	C	C	C	A	+	B	A	A					1	0	1	5	1	1	15.9

HIGH AND ELEMENTARY SCHOOL GRADES

EXPRESSED AS DEVIATIONS FROM MEAN

Individual No.	High School 1st year			7th Gr.			6th Gr.			5th Gr.			4th Gr.		
	F_M	F_E	F_{AV}	7_M	7_E	7_H	6_M	6_E	6_H	5_M	5_E	5_H	4_M	4_E	4_H
4	3	−1	0	0	3	−1	1	1	0	1	2	1	0	−1	1
9	−10	−10	−5	1	2	0	1	3	1	1	3	1	0	1	1
10	1	−1	1	−2	−3	0	−1	0	0	1	−1	1	0	4	−0
12	3	−4	3	0	0	−1	−4	0	−1	−3	−3	−2	1	2	1
14	−4	−6	−5	−2	3	2	0	7	1	0	6	4	−2	8	−3
15	10	8	9	−3	−4	−4	−2	−6	−3	0	−3	0	−2	−4	−2
16	−1	4	3	1	−3	−2	2	−3	0	1	−2	0	0	−3	−1
17	1	−2	−2	0	1	1	2	3	2	1	3	1	1	3	3
18	−3	−1	−2	1	2	1	3	5	2	0	0	0	0	2	−2
24	10	8	7	0	−1	−3	−4	−6	−2	0	−5	−2	−4	−6	−3
25	7	4	5	0	3	−3	−1	3	2	−3	3	−2	−4	−3	−3
29	7	4	5	−3	−2	−2	−3	−4	−2	−1	2	0	−3	−4	−1
36	3	4	0	−3	−3	−2	−2	3	1	0	3	2	−4	2	−2
38	3	8	4	−2	−5	−1	−3	−5	0	2	−3	1	−2	−4	0
39	5	0	−1	0	1	−2	0	3	−2	1	2	1	1	2	−1
40	−6	4	0	0	0	1	1	−1	0	−4	−5	−1	−4	−5	4
41	0	−1	−3	2	0	−1	1	−2	0	0	3	2	2	0	1
42	−1	−10	−6	2	4	0	−1	3	1	−3	2	1	−1	3	−3
45	7	1	1	2	−1	−1	−3	0	−2	−3	−3	−2	−4	−6	−3
46	−4	−2	−3	1	−1	1	−2	−1	0	1	0	1	−2	1	−3
47	5	1	2	0	2	0	−1	−1	0	0	3	0	−1	2	1
53	5	4	6	−3	−3	−3	−3	−3	−2	−1	0	−2	−2	−3	−3
65		−2	0	2	0	1	2	2	−1	−1	0	1	3	2	1
76		−6	−5	1	1	0	1	1	1	−3	−2	−1	−1	3	1
89		−2	−3	1	2	0	2	−3	1	3	−3	1	1	0	1
98		−1	−2	2	3	2	4	2	2	1	3	3	5	5	4
103	−4	3	−1	2	−1	0	1	−2	−3	0	−2	−2	2	−3	−1
106	−3	−6	−3	0	0	0	0	1	0	0	−2	1	2	−2	1
109	2	0	0	0	3	−1	−1	−1	0	1	−2	1	−1	−3	1
110	12	9	10	−1	−2	−3	−3	−3	−2	1	−6	−3	−1	−5	−3
112	−3	−2	−2	1	−2	0	1	2	0	1	0	1	−3	−6	−1
113	4	−2	1	0	6	0	0	5	0	0	6	2	2	2	1
116	4	−2	2	0	−2	−1	1	−1	0	−3	0	−2	−1	−4	−1
123	3	−4	−1	0	6	1	0	4	1	0	3	2	−3	−3	1
125	−5	−4	−4	1	3	0	0	2	−1	−3	1	−2	3	2	1
127	0	−4	−2	−1	2	0	0	3	−2	−3	0	−1	−2	0	1
128	5	−2	0	0	2	−1	−1	0	−1	−1	0	0	−2	0	1
129	8	9	9	−1	−3	−4	0	−5	−2	−3	−6	−3	−1	−4	−1
130	−2	−2	−3	0	2	0	1	2	0	2	1	2	−1	0	−2
135	−1	5	2	−1	3	−1	1	2	0	0	−2	1	−1	−1	−1
138	8	6	6	−2	−1	−1	−4	0	−4	−2	−4	−2	−3	−1	−1
143	6	7	7	−1	−2	−2	−1	−3	−3	−3	−5	−3	−3	−5	−3
145	2	5	3	−1	−1	0	0	0	0	1	−5	0	−3	−5	−1
147	6	0	3	0	3	−1	0	−4	−1	−3	−1	2	−2	−4	−3
148	2	5	3	−2	−2	−1	0	−1	0	−2	−5	−2	−1	−3	−1
151	5	−2	2	0	−2	0	0	0	−1	−1	6	−1	0	2	−1
152	0	1	1	1	2	0	0	0	0	−1	−4	−1	−1	0	−1
153	−1	2	1	−1	2	0	0	1	−1	−3	2	−3	−3	−1	1
154	0	5	3	2	−2	0	3	0	1	1	−6	−1	1	−1	1
156	−2	2	−1	−1	1	0	1	1	0	−1	−6	−1	0	−6	−1
157	4	−2	0	4	−2	1	1	2	−2	0	2	−3	1	−2	1
160	0	1	0	1	−1	0	−1	−3	−1	−3	−4	1	−1	1	1
162	−1	0	−2	0	0	0	0	2	2	0	−2	3	2	−1	2
166	6	0	3	1	1	0	0	2	0	0	−1	1	1	−2	1
168	6	3	4	0	0	1	0	−3	−1	−1	−2	1	−3	−3	1
169	−8	3	−1	1	0	2	1	−3	0	2	−6	−2	2	−6	1
172	−2	−8	−5	0	3	1	0	2	1	−2	−2	2	0	−1	1
173	2	2	1	0	1	−1	−1	1	−1	1	0	0	1	−3	0
174	−5	7	5	0	−4	−4	0	−4	−4	0	−6	−3	1	−4	1